The Learning Effect of Online Discussions: New Active Form of Learning under COVID-19

Jaeseo Lim
   Seoul National University

Hyunwoong Ko
   Seoul National University

Jooyong Park
   Seoul National University

Jungjoon Ihm (✉ ijj127@snu.ac.kr)
   Seoul National University

Abstract

Background: COVID-19 created confusion in all areas of society and is no exception in education. In line with the demands of the times, we investigated the applicability of active learning online for the health professional.

Methods: This study was conducted in the school of dentistry, a selective university. Eighty-two dentistry students participated in this study in the second semester of 2020. The students were randomly assigned to four different groups and learned according to given experimental procedures. Their performance in the final test scores was analyzed using ANOVAs. Subsequently, we systematically analyzed past offline active learning and online experiments.

Results: As with past offline experiments, self-study before discussion groups have achieved high learning outcomes in verbatim type and transfer type items. Online self-study groups scored slightly lower than offline self-study groups, but online self-study groups scored higher than offline and online lecture groups.

Conclusions: Our findings denote that active learning for health professionals was applicable online. When self-study precedes the discussion, it becomes richer, thereby increasing students' learning outcomes.

Background

The novel coronavirus infection (also known as COVID-19), which began in December 2019, has become a global pandemic. Though the situation seems to be improving for some countries and regions, the effects have been massive. COVID-19 has impacted all areas of society, and education is no exception. In February 2020 in South Korea, face-to-face classes were banned in elementary, middle, high schools and universities, and "online" schools were announced in March [1, 2, 3]. Remote education was to be implemented for 3 million students as a new national educational system. Likewise, in Hong Kong 1,302 schools closed in-person learning, and 1 million students stayed home [4].

Many countries have opted for online education as a result of the pandemic. Online education refers to a form of learning that takes place regardless of spatial or temporal restrictions by using digital tools such as the Internet, smartphones, and other mobile devices [4]. While online learning differs slightly for scholar to scholar in concept and definition, it exists in various types, such as web-based learning, e-learning, cyber learning, distance learning, or mobile learning [5, 6]. Some teachers, schools, and institutes upload their lectures online for students to access and use Google Classrooms, WebQuest, and other online learning platforms [7, 8].

In response to COVID-19, even education for health professions, including medical education, has quickly transitioned its entire curriculum to online formats, including subjects such as basic science and behavioral science [9]. Educators and educational institutions for those pursuing health professions have

made changes in response to contemporary needs and have attempted to create better, more effective learning methods for their students. For decades, these educators have been working to transform their pedagogy by reducing or even eliminating lectures [9, 10] and by implementing team-based, active, and self-directed learning [11, 12]. Within the past few years, medical educators have been converting classroom-style learning to one-way lectures and providing instruction for asynchronous learning [13]. Similarly, an alternative to lectures is having discussions, a student-centered learning method. Evidence that discussion facilitates learning was confirmed in experimental studies comparing the effectiveness of learning after watching one-way lectures and after watching discussion-style videos online [14, 15]. The results of this experiment showed that students who watched the discussion-style videos, and those who watched them while discussing the learning materials, had greater performance. However, many students majoring in health professions still study in monologue-style lecture sessions, laboratory sessions, and simulations [9].

In addition, there has been a lack of quantitative and empirical research on students' performance in online education under COVID-19 from the medical field, including medical education and other fields related to health science (e.g., dentistry) [16, 17, 18]. It is especially rare to find a study that evaluates student achievement by conducting active learning online. In this study, we conducted a study online that provided active learning for health professionals before the spread of COVID-19 [10]. In previous research, Lim and colleagues explored the ICAP framework, in which active pre-learning activities should be preceded by the discussion, and whether the framework can be utilized more effectively to train health professionals [10, 19, 20]. They proved the effectiveness of discussion on student's learning within health-related subjects including medicine, dentistry, veterinary medicine, and nursing and concluded when discussion was most effective. Therefore, based on the study of desirable active learning in health professions, while effectively replicating past experiments, we seek to investigate the direction of active learning online that ensures the quality of learner-centered classes that can be effectively applied in future pandemics and other crises.

Methods

For this experiment, we predetermined sample sizes of at least 19 participants per group in order to compare with the offline experiment before COVID-19 [10]. In addition, we conducted power analyses using the G*Power to determine whether the design in this experiment had enough power [21]. Groups of 19 participants provided 80% power to detect an effect size of 0.87 (alpha =.05).

Participants

Participants (N = 82) were recruited from school of dentistry at Seoul National University for this study. The study was reviewed and approved by the Institutional Review Board (IRB) of Seoul National University School of Dentistry (approval No. S-D20200053).

Survey on Prior Knowledge

We used a survey to check participants' prior knowledge or interest on the stimulus topic in order to minimize the effect of prior knowledge on the experiment [22, 23]. Six survey items were assessed using a 7-point Likert scale ranging from "do not know about it at all (1)" to "know very well about it (7)." The survey included six items, two of which were related to the topic of the learning material (the criminal procedure code and the accusation and charge) and four that were unrelated (the genome project, civil law, the legalization of same-sex marriage, and the Special Acton Sexual Violence).

Video Lecture and Learning Material

The video lecture used in this study showcased a monologue-style lecture. The lecture was on the topic of law, particularly concerning the criminal procedure code, accusation, and charge. This topic was chosen because it is not typically taught in undergraduate courses. This is an especially unfamiliar topic for students studying health. This topic was also chosen because its assessments include a finite set of answers for transfer type items used to evaluate students' learning outcomes. The video lecture was 18 minutes in length.

Learning materials derived from the contents of the video lecture were seven pages long. These materials also covered information on the same topic: the criminal procedure code, accusation, and charge.

Final Test Items

Items for the final test were comprised of verbatim type and transfer type items. Verbatim type items were comprised of 10 questions with answers that were directly inferable from the learning materials. These questions consisted of short answer or multiple choice questions, with a total of 10 questions worth one point each. For example, one question states, "Given that there is no one who filed an accusation against a crime subject for prosecution, prosecutors must designate a person who can file the complaint within ( ) days upon the request of the stakeholders." Students who filled in a correct answer would have written "10" in the blank space given.

Transfer type items consisted of four questions, each requiring total comprehension of the given materials and application of these materials to new situations. These items were 15 points each; thus, a perfect score on these test items would total 25 points. An example of these questions is as follows: "The under-aged victim (V) accused the defendant (D) of rape, but eventually withdrew her accusation on February 1st, 2017. Afterwards, V's father (F), legal representative of V, accused D on February 10th, 2017. D was charged with rape and was convicted of the crime at the first trial. However, D made an appeal claiming that F's complaint should not take effect since V had already withdrawn her complaint, and therefore, the prosecutor's indictment was against the provisions of the law. Will the Court of Appeals accept D's claim?"

Procedure

This experiment was designed to replicate and build off of the study conducted in the past that used an offline lecture and set of learning materials [10]. At the same time, this experiment was designed to determine whether the same extent of active learning is possible through online discussion. Thus, the overall experimental methods and procedures were largely the same, excluding one variable that posed questions on the learning processes, which did not prove effective in the past experiment. The use of worksheets alongside instruction by the lecturer or professor was added as a new variable.

All students participated in this experiment through Zoom, an online meeting platform. Participants were instructed to fill out the background knowledge survey to determine their level of prior knowledge and interest. Participants were randomly assigned to groups of two different study conditions: lecture and self-study. Through the breakout rooms function within Zoom, students were randomly assigned to groups of three or four. In the lecture groups, participants were instructed to listen to a video lecture while looking at the learning materials for 18 minutes. In the self-study group, participants were instructed to study written materials by themselves for 18 minutes. Subsequently, participants from each of the two categories were divided into four different groups of three or four participants prior to discussions. Worksheets were provided for some groups during discussion, and not provided for others. The discussion entailed participants talking freely with peers about what they had studied for 20 minutes. Two of the groups received worksheets from professors to discuss with their peers. Finally, students were given 20 minutes to complete a final test.

To sum up, depending on our experimental design, students were randomly assigned to each of the four groups: two study conditions (lecture versus self-study) and two intervention conditions (discussion with worksheet versus free discussions). Specifically, there were four groups: (1) the lecture and discussion (LD) group, (2) the lecture and discussion with worksheet (LW) group, (3) the self-study and discussion (SD) group, and (4) the self-study and discussion with worksheet (SW) group.

Data Analyses

Data analysis was carried out using SPSS version 23 (SPSS Inc., Chicago, USA). The data was analyzed using descriptive statistics (means, standard deviations (SDs), minimums, maximums). One-way or two-way analysis of variance (ANOVA) was conducted to compare each of four groups. P value < 0.05 was considered to indicate a statistically significant difference.

Results

Learning Outcomes of Active Learning Online

Regardless of which group students were in, all students completed the given prior knowledge survey. The learning materials used in this experiment were law-related and deal with accusation, charge, and recognition of criminal procedure code. First, we analyzed survey items by topic: criminal procedure code (M = 1.98, SD = 1.06), accusation and charge (M = 2.45, SD = 1.07), genome project (M = 3.73, SD = 1.93), civil law (M = 1.86, SD = 0.99), legalization of same-sex marriage (M = 2.47, SD = 1.06) and Special Act on

Sexual Violence (M = 2.64, SD = 1.47). Among these items, scores of topic-relevant items (criminal procedure code and accusation and charge) were significantly lower than that of the genome project, which was the highest, $t(164)$ = 7.29, P < 0.001, $t(164)$ = 5.34, P < 0.001, respectively. There was also no difference between the four groups in their background knowledge of the learning material.

The mean values and standard deviations of the scores for total, verbatim type, and transfer type items are presented in **Table 1**. In order to rule out the subjectivity of scoring as much as possible, 15 final test papers were randomly selected and were marked by two other raters (20% of the total). For all types of question items, the agreement measured by Intraclass correlation was 0.89 for the analysis {ICC (3.k)}. Accordingly, the agreement measured by ICC was good, so the remaining final test paper was marked by the first author.

First, analyses were performed to compare the main effect of the SD condition, as well as the interaction effect between these two conditions on learning outcomes, particularly transfer type items. The main effect of the study condition (lecture versus self-study) was significant, $F(1, 78)$ = 23.47, P < 0.001, partial eta-squared =.231. However, the main effect of the discussion condition (discussions with worksheet versus free discussions) had no significant difference, $F(1, 78)$ = 2.40, P = 0.125, partial eta-squared =.030. The interaction effect of these two factors was significant, $F(1, 78)$ = 9.04, P < 0.001, partial eta-squared =.104. (**Fig. 1**)

Subsequently, significant differences were found among the groups in total score ($F(3, 78)$ = 12.35, P < 0.001, partial eta-squared =.322), verbatim type items ($F(3, 78)$ = 5.22, P = 0.002, partial eta-squared =.167), and transfer type items ($F(3, 78)$ = 12.21, P < 0.001, partial eta-squared =.320).

We conducted a planned comparison to find out any difference between groups in accordance with our hypotheses. First of all, the SD group was significantly higher than the SW group, the LD group, and the LW group, $t(78)$ = 3.42, P = 0.001, d = 1.299; $t(78)$ = 5.18, P < 0.001, d = 1.747; $t(78)$ = 5.28, P < 0.001, d = 1.568, respectively. However, there were no significant differences between the LD group and the LW group, between the LW group and the SW group, and between the LD group and the SW group, $t(78)$ = 0.23, P = 0.817, d = 0.063; $t(78)$ = 1.87, P = 0.066, d = 0.542; $t(78)$ = 1.68, P = 0.097, d = 0.545, respectively.

For verbatim type items, the SD group had significantly higher scores than those of the SW group, the LD group, and the LW group, $t(78)$ = 2.37, P = 0.020, d = 0.995; $t(78)$ = 2.19, P = 0.032, d = 0.699; $t(78)$ = 3.90, P < 0.001, d = 1.298, respectively. However, there were no significant differences between the LD group and the LW group, between the LW group and the SW group, and between the LD group and the SW group, $t(78)$ = 1.75, P = 0.084, d = 0.457; $t(78)$ = 1.52, P = 0.133, d = 0.460; $t(78)$ = 0.22, P = 0.829, d = 0.065.

For transfer type items, the SD group was significantly higher than the SW group, the LD group, and the LW group, $t(78)$ = 3.26, P = 0.002, d = 1.088; $t(78)$ = 5.69, P < 0.001, d = 1.828; $t(78)$ = 4.04, P < 0.001, d = 1.338, respectively. The SW group had significantly higher scores than those of the LD group, $t(78)$ = 2.33,

P = 0.022, d = 0.770, while there were no significant differences between the LD and the LW groups, and between the LW and the SW groups, $t(78)$ = 1.02, P = 0.312, d = 0.300; $t(78)$ = 1.27, P = 0.208, d = 0.386.

Comparison of Active Learning Offline and Online

We intended to replicate the results of the past experiment [10] and at the same time determine whether the same extent of active learning is possible through online discussion. Thus, we investigated whether the online discussion improved the student's learning outcomes, especially the greater the transfer.

The mean values and standard deviations of all groups are provided in **Fig. 2**. We compared offline study [10] with online study, separating groups according to lecture or self-study, which were the learning activities prior to the discussion. The four groups were lecture and discussion offline (LD Offline, n = 50), lecture and discussion online (LD Online, n = 40), self-study and discussion offline (SD Offline, n = 58), and self-study and discussion online (SD Online, n = 42). The four groups had significant differences in total scores, verbatim types items and transfer type items, $F(3, 186)$ = 23.71, P < 0.001, partial eta-squared =.277; $F(3, 186)$ = 10.49, P < 0.001, partial eta-squared =.145; $F(3, 186)$ = 19.82, P < 0.001, partial eta-squared =.242.

We carried out a post hoc analysis using the Tukey test to compare the four groups. The total mean values of the SD Offline group (M = 16.57, SD = 3.70) was the highest of the four groups, followed by the SD Online group (M = 14.31, SD = 4.03). Ps <.01. However, the SD Offline group had no significant difference with the LD Online group (M = 12.78, SD = 3.14), P =.235. In addition, there were significant differences between the SD Offline group, the SD Online group, and the LD Online group (M = 10.08, SD = 4.67) in order, Ps <.05.

For verbatim type items, the SD Offline group (M = 8.57, SD = 1.40) was the highest among the four groups, Ps <.05. The SD Online group (M = 7.67, SD = 1.82) was significantly higher than the LD Online group (M = 6.48, SD = 2.69), P = 0.021, but there was no significant difference from the LD Offline group (M = 7.36, SD = 1.50), P = 0.005. There was no significant difference between the LD Offline group and the LD Online group, P = 0.114.

For transfer type items, the SD Offline group (M = 8.00, SD = 3.12) was significantly higher than the LD Offline group (M = 5.42, SD = 2.29) and the LD Online group (M = 3.60, SD = 3.05), Ps < 0.001, while there was no significant difference from the SD Online group (M = 6.64, SD = 3.02), P = 0.097. The SD Online group had no significant difference from the LD Offline group, P = 0.183, and the LD Offline group was significantly higher than the LD Online group, P = 0.017. (**Table 2**)

Learning Outcomes of Active Learning Online

Regardless of which group students were in, all students completed the given prior knowledge survey. The learning materials used in this experiment were law-related and deal with accusation, charge, and recognition of criminal procedure code. First, we analyzed survey items by topic: criminal procedure code (M = 1.98, SD = 1.06), accusation and charge (M = 2.45, SD = 1.07), genome project (M = 3.73, SD = 1.93),

civil law (M = 1.86, SD = 0.99), legalization of same-sex marriage (M = 2.47, SD = 1.06) and Special Act on Sexual Violence (M = 2.64, SD = 1.47). Among these items, scores of topic-relevant items (criminal procedure code and accusation and charge) were significantly lower than that of the genome project, which was the highest, $t(164)$ = 7.29, P < 0.001, $t(164)$ = 5.34, P < 0.001, respectively. There was also no difference between the four groups in their background knowledge of the learning material.

The mean values and standard deviations of the scores for total, verbatim type, and transfer type items are presented in **Table 1**. In order to rule out the subjectivity of scoring as much as possible, 15 final test papers were randomly selected and were marked by two other raters (20% of the total). For all types of question items, the agreement measured by Intraclass correlation was 0.89 for the analysis {ICC (3.k)}. Accordingly, the agreement measured by ICC was good, so the remaining final test paper was marked by the first author.

First, analyses were performed to compare the main effect of the SD condition, as well as the interaction effect between these two conditions on learning outcomes, particularly transfer type items. The main effect of the study condition (lecture versus self-study) was significant, $F(1, 78)$ = 23.47, P < 0.001, partial eta-squared =.231. However, the main effect of the discussion condition (discussions with worksheet versus free discussions) had no significant difference, $F(1, 78)$ = 2.40, P = 0.125, partial eta-squared =.030. The interaction effect of these two factors was significant, $F(1, 78)$ = 9.04, P < 0.001, partial eta-squared =.104. (**Fig. 1**)

Subsequently, significant differences were found among the groups in total score ($F(3, 78)$ = 12.35, P < 0.001, partial eta-squared =.322), verbatim type items ($F(3, 78)$ = 5.22, P = 0.002, partial eta-squared =.167), and transfer type items ($F(3, 78)$ = 12.21, P < 0.001, partial eta-squared =.320).

We conducted a planned comparison to find out any difference between groups in accordance with our hypotheses. First of all, the SD group was significantly higher than the SW group, the LD group, and the LW group, $t(78)$ = 3.42, P = 0.001, d = 1.299; $t(78)$ = 5.18, P < 0.001, d = 1.747; $t(78)$ = 5.28, P < 0.001, d = 1.568, respectively. However, there were no significant differences between the LD group and the LW group, between the LW group and the SW group, and between the LD group and the SW group, $t(78)$ = 0.23, P = 0.817, d = 0.063; $t(78)$ = 1.87, P = 0.066, d = 0.542; $t(78)$ = 1.68, P = 0.097, d = 0.545, respectively.

For verbatim type items, the SD group had significantly higher scores than those of the SW group, the LD group, and the LW group, $t(78)$ = 2.37, P = 0.020, d = 0.995; $t(78)$ = 2.19, P = 0.032, d = 0.699; $t(78)$ = 3.90, P < 0.001, d = 1.298, respectively. However, there were no significant differences between the LD group and the LW group, between the LW group and the SW group, and between the LD group and the SW group, $t(78)$ = 1.75, P = 0.084, d = 0.457; $t(78)$ = 1.52, P = 0.133, d = 0.460; $t(78)$ = 0.22, P = 0.829, d = 0.065.

For transfer type items, the SD group was significantly higher than the SW group, the LD group, and the LW group, $t(78)$ = 3.26, P = 0.002, d = 1.088; $t(78)$ = 5.69, P < 0.001, d = 1.828; $t(78)$ = 4.04, P < 0.001, d = 1.338, respectively. The SW group had significantly higher scores than those of the LD group, $t(78)$ = 2.33,

P = 0.022, *d* = 0.770, while there were no significant differences between the LD and the LW groups, and between the LW and the SW groups, *t*(78) = 1.02, *P* = 0.312, *d* = 0.300; *t*(78) = 1.27, *P* = 0.208, *d* = 0.386.

Comparison of Active Learning Offline and Online

We intended to replicate the results of the past experiment [10] and at the same time determine whether the same extent of active learning is possible through online discussion. Thus, we investigated whether the online discussion improved the student's learning outcomes, especially the greater the transfer.

The mean values and standard deviations of all groups are provided in **Fig. 2**. We compared offline study [10] with online study, separating groups according to lecture or self-study, which were the learning activities prior to the discussion. The four groups were lecture and discussion offline (LD Offline, *n* = 50), lecture and discussion online (LD Online, *n* = 40), self-study and discussion offline (SD Offline, *n* = 58), and self-study and discussion online (SD Online, *n* = 42). The four groups had significant differences in total scores, verbatim types items and transfer type items, $F(3, 186) = 23.71$, $P < 0.001$, partial eta-squared =.277; $F(3, 186) = 10.49$, $P < 0.001$, partial eta-squared =.145; $F(3, 186) = 19.82$, $P < 0.001$, partial eta-squared =.242.

We carried out a post hoc analysis using the Tukey test to compare the four groups. The total mean values of the SD Offline group (*M* = 16.57, *SD* = 3.70) was the highest of the four groups, followed by the SD Online group (*M* = 14.31, *SD* = 4.03). *Ps* <.01. However, the SD Offline group had no significant difference with the LD Online group (*M* = 12.78, *SD* = 3.14), *P* =.235. In addition, there were significant differences between the SD Offline group, the SD Online group, and the LD Online group (*M* = 10.08, *SD* = 4.67) in order, *Ps* <.05.

For verbatim type items, the SD Offline group (*M* = 8.57, *SD* = 1.40) was the highest among the four groups, *Ps* <.05. The SD Online group (*M* = 7.67, *SD* = 1.82) was significantly higher than the LD Online group (*M* = 6.48, *SD* = 2.69), *P* = 0.021, but there was no significant difference from the LD Offline group (*M* = 7.36, *SD* = 1.50), *P* = 0.005. There was no significant difference between the LD Offline group and the LD Online group, *P* = 0.114.

For transfer type items, the SD Offline group (*M* = 8.00, *SD* = 3.12) was significantly higher than the LD Offline group (*M* = 5.42, *SD* = 2.29) and the LD Online group (*M* = 3.60, *SD* = 3.05), *Ps* < 0.001, while there was no significant difference from the SD Online group (*M* = 6.64, *SD* = 3.02), *P* = 0.097. The SD Online group had no significant difference from the LD Offline group, *P* = 0.183, and the LD Offline group was significantly higher than the LD Online group, *P* = 0.017. (**Table 2**)

Discussion

Although there has been a focused attention on the need to care for patients and communities, there has not been sufficient focus to address the effect of the advent of COVID-19 on medical education. COVID-19 has effected educational institutions worldwide; in the country alone, thousands of schools and

universities faced closure in given the limited time span to enforce social distancing [7]. While in the midst of the pandemic crisis, it is crucial to see that this situation offers the opportunity to rethink modern educational content, purposes, and systems. The profound effects of this disease call for changes in modern education.

Amidst the pandemic, as an alternative to existing teaching methods many countries around the world have adopted online learning such as web-based learning and e-learning. Although online learning has existed before, it has become more sophisticated and diverse through the development of information technology and the needs of the times. Accordingly, health-related education including medical education began to change existing learning methods to include online formats. In addition, educators tried to minimize or eliminate one-way lectures and improve student-centered learning methods [9, 10]. As a good alternative to these approaches, Lim and colleagues proposed an active learning method for health professionals [10]. They confirmed the effectiveness of discussions in learning, while exploring which preliminary activities before the discussions were most effective. The experiment showed that the more active the students were in preliminary activities before discussions, the greater the effect of the discussion, which was shown to be the students' high learning performance. Thus, the present study was conducted to add additional evidence on the effectiveness of discussions with dentistry students by replicating the previous study and to investigate whether active learning could be applied online.

The results of this study obtained similar results as those of the past [10]. The main effect of the study condition (lecture versus self-study) was significant while the main effect of the discussion condition had no significance (discussions with worksheet versus free discussions). However, the interaction effect of these two factors was significant. In detail, the SD group had the highest performance when comparing total scores, for both verbatim and transfer type items of the final test. The other three groups had no significant differences, but the LW group showed the lowest score. For scores on verbatim type items, the SD group had the highest scores among the four groups. It was not significant, but after the SD group the scores were highest in order of the LD group, the SW group, and the LW group. For transfer type items, the scores were highest in the order of the SD group, the SW group, and the two lecture groups, with no significant differences between the two lecture groups. These difference in scores and patterns between the groups were similar for all types of items. Thus, students who studied by themselves before the discussions had better outcomes compared to those who listened to a lecture. This suggests that student participation in more active learning activities rather than simply listening to lectures enhances the effect of discussions, which leads to higher learning outcomes.

In addition, structured discussions, such as those including the instructors' intervention (e.g., a worksheet), did not work well. The LW group that received the worksheet had a lower score than the LD group that did not, and the SW group that received the worksheet had a lower score than the SD group who did not. Interestingly, in the two groups who listened to the lecture, the instructors' intervention had some effect on improving the students' scores, but in the two groups who conducted the self-study, the instructors' intervention actually lowered the students' scores. These results suggest that free

discussions, such as solving each other's questions or discussing topics in more depth, are effective discussions for self-study groups. On the other hand, the lecture groups in which students participated in more passive learning seemed to allow for deeper learning through the intervention of instructors. These are results that support past studies that show that one-way lectures cannot induce students to learn as deeply as other methods do [24, 25, 26].

Next, we compared the offline experiment [10] with the current online experiment. As a result of the analysis, the SD Offline group was the highest of the four groups in total scores, followed by the SD Offline group. The LD Offline group had no significant difference from the LD Online group, and the LD Online group had the lowest score. For verbatim type items, the SD Offline group had the highest score followed by the SD Online group, and there was no significant difference between the LD Offline group and the LD Online group. Lastly, the SD Offline group had the highest score for transfer type items, and the LD Online group had the lowest score. In summary, regardless of whether learning was offline or online, the group that had discussions and that was preceded by self-study scored higher. The important point is that in these groups, students participate in the learning process more actively and try to understand the learning materials themselves. However, online groups generally scored lower than offline groups. This is because students were not only awkward with the online platform, but were also unfamiliar with the progress of active learning online. In the future, if students experience online learning, results are expected to be similar to those of offline learning. Nevertheless, it is important that patterns in the results of offline learning experiments and online learning experiments were similar.

The main results of our experiment support the ICAP framework [19, 20]. In parallel with the framework that the engagement of students promotes greater learning outcomes, the self-study groups (e.g., the SD, SW, SD Offline, and SD Online groups) scored much higher than the lecture groups (e.g., the LD, LW, LD Offline, and LD Online groups). That is, the results are consistent with Chi and colleagues' ICAP framework showing the learning benefits of active learning. Moreover, based on the results of this study, active learning online, including discussions, is expected to be applied to students with health-related majors, including medical and dentistry majors (e.g., Team-based learning online, Project-based learning online). However, future studies are needed to expand and generalize our findings. One limitation of the present study is that the experimental group was limited to dentistry students. Thus, future studies need to recruit students from other health professions (e.g., medical, veterinary etc.) to conduct comparative studies. The second limitation is that past offline experiments and current experiments did not entirely follow the same process. Offline experiments [10] used a question-generation condition, but we added a new variable called instructors' intervention (worksheets) during the discussion. Since the question-generation variable did not have a significant main effect in past experiments, the new variable of instructors' intervention during discussion is meaningful in that we explored new factors in active learning. Moreover, careful caution is needed since the intervention in a discussion may be good or bad depending on the situation.

This study explores effective learning methods for health professionals. Students of health professions have thus far pursued large amounts of knowledge, but they require effective learning

methods to better remember, apply, or utilize this knowledge in new situations [27, 28, 29]. In this regard, this study presents a learning method that can improve their learning, which is a discussion after self-study. More importantly, in accordance with the demands of the times of COVID-19, this study explored a new form of active learning online and assessed whether the active learning method of discussion can be applied online. Possible follow-up studies centered on some of the above discussions, as well as active learning studies online, will create new changes in future education.

Conclusion

The pandemic changes the practical and logistical schemes of educations for health professions. In response to COVID-19, online or non-face-to-face learning methods are required. Thus, our experiment is timely and can be appropriate to cope with these times in the context of health education. Our findings support the ICAP framework and provide practical implications for education for health professionals. It is better for students to study by themselves before discussions than to listen to lectures offline or online. In particular, when students participate actively in learning activities before discussions, they increase their transfer of learning. Thus, through new approaches such as developing and applying existing methods online, we look forward to creating more effective and diverse educational methods for health professionals.

Abbreviations

lecture and discussion (LD) group; lecture and discussions with worksheet (LW) group; self-study and discussion (SD) group; self-study and discussions with worksheet (SW) group; lecture and discussion offline (LD Offline) group, lecture and discussion online (LD Online) group, self-study and discussion offline (SD Offline) group, and self-study and discussion online (SD Online) group.

Declarations

Acknowledgments

Authors' contributions

All authors have read and approved the manuscript. Conceptualization: LJ; Methodology: LJ; Formal analysis: LJ; Data curation: LJ; Investigation: LJ, KH, PJ, IJ; Writing - original draft preparation: LJ, KH; Writing - review and editing: PJ, IJ

Funding

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The study was approved by the Institutional Review Board (IRB) of Seoul National University School of Dentistry (IRB No. S-D20200053). All the experiment protocol for involving human participants was in accordance to guidelines of national/international/institutional ethical standards or Declaration of Helsinki. All participants were aware that they were taking part in this research and gave informed consent in addition to confirming that they would allow us to use their collected data anonymously for publication. All the data were anonymously collected and analyzed.

Consent for publication

Not Applicable.

Competing Interests

The authors declare that they have no competing interests.

References

[1] Ministry of Education of Korea. Press Releases. Year of Publication. http://english.moe.go.kr/main.do?s=english.

[2] BBC News Services. Coronavirus: How South Korea is teaching empty classrooms. 2020. https://www.bbc.com/news/world-asia-52230371.

[3] The Korea Herald. The devilish coronavirus. 2020. http://www.koreaherald.com/view.php?ud=20200520000156&ACE_SEARCH=1.

[4] Clark D. Psychological myths in e-learning. Med Teach. 2002;24(6):598–604.

[5] Jacobson MJ. Cognitive visualisations and the design of learning technologies. Int. J. Learn. Technol. 2004;1(1):40–62.

[6] Khan BH. Managing e-learning: Design, delivery, implementation, and evaluation. Edition. Publisher Location: IGI Global; 2005.

[7] Toquero CM. Challenges and opportunities for higher education amid the COVID-19 pandemic: The Philippine context. Pedagog Res. 2020;5(4).

[8] Fox R. SARS epidemic: Teachers' experiences using ICTs. In *Beyond the Comfort Zone: Proceedings of the 21st ASCILITE Conference*:319–327Citeseer 2004.

[9] Rose S. Medical student education in the time of COVID-19. JAMA. 2020;323(21): 2131-2132.

[10] Lim J, Ko H, Yang JW, Kim S, Lee S, Chun MS, Ihm J, Park J.. Active learning through discussion: ICAP framework for education in health professions. BMC Med Educ. 2019;19(1), 1-8.

[11] Irby DM, Cooke M, O'Brien BC. Calls for reform of medical education by the Carnegie Foundation for the Advancement of Teaching: 1910 and 2010. Acad Med. 2010;85(2):220–227.

[12] Skochelak SE, Stack SJ. Creating the medical schools of the future. Acad Med. 2017;92(1):16–19.

[13] Tucker B. The flipped classroom. Educ Next. 2012;12(1):82–83.

[14] Daradoumis T, Bassi R, Xhafa F, Caball′e S. A review on massive e-learning (MOOC) design, delivery and assessment. In *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*:208–213IEEE 2013.

[15] Hew KF, Cheung WS. Students' and instructors' use of massive open online courses(MOOCs): Motivations and challenges. Educ Res Rev. 2014;12:45–58.

[16] Mujayanto R, Indraswary R. Covid-19 pandemic and challenges of dentistry: Differential diagnosis of COVID-19 Enanthema. Eur J Dent. 2020;14(Suppl1):S179.

[17] Meng L, Hua F, Bian Z. Coronavirus disease 2019 (COVID-19): emerging and future challenges for dental and oral medicine. J Dent Res. 2020;99(5):481–487.

[18] Usak M, Masalimova AR, Cherdymova EI, Shaidullina AR. New playmaker in science education: Covid-19. J Baltic Sci Educ. 2020;19(2):180.

[19] Chi MT. Active-constructive-interactive: a conceptual framework for differentiating learning activities. Top Cogn Sci. 2009;1(1):73–105.

[20] Chi MT, Wylie R. The ICAP framework: linking cognitive engagement to active learning outcomes. Educ Psychol. 2014;49(4):219–43.

[21] Faul F, Erdfelder E, Lang AG, Buchner A. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39(2):175–191.

[22] Beyer BK. Practical strategies for the teaching of thinking. Edition. Publisher Location: ERIC; 1987.

[23] Miyake N, Norman DA. To ask a question, one must know enough to know what is not known. J Verb Learn Verb Behav. 1979;18(3):357–364.

[24] Hrepic Z, Zollman DA, Rebello NS. Comparing students' and experts' understanding of the content of a lecture. J Sci Educ Technol. 2007;16(3):213–224.

[25] Wieman C, Perkins K. Transforming physics education. Phys Today. 2005;58(11):36.

[26] Poh MZ, Swenson NC, Picard RW. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. IEEE Trans Bio Med Eng. 2010;57(5):1243–1252.

[27] Moon SH, Myung SJ, Yoon HB, Park JB, Kim JW, Park WB. Deliberate practice as an effective remediation strategy for underperforming medical students focused on clinical skills: a prospective longitudinal study. J Korean Med Sci. 2019;34(11).

[28] Graffam B. Active learning in medical education: strategies for beginning implementation. Med Teach. 2007;29(1):38–42.

[29] Kim S, Yang, JW, Lim J, Lee S, Ihm J, & Park J. The impact of writing on academic performance for medical students. BMC Med Educ. 2021;21(1), 1-8.

Tables

Table 1. Learning outcomes comparison for types of items and study conditions.

Type of items	LD ($n = 21$)	LW ($n = 19$)	SD ($n = 22$)	SW ($n = 20$)	P
Total score (25 points)	10.19 (4.19)	9.89 (5.21)	16.55 (2.99)	12.30 (4.80)	<0.001
Verbatim (10 points)	7.05 (2.67)	5.84 (2.64)	8.50 (1.22)	6.90 (1.92)	0.002
Transfer (15 points)	3.14 (2.74)	4.05 (3.31)	8.05 (2.62)	5.20 (2.61)	<0.001

Note. Data are shown as mean values (standard deviations). LD = lecture and discussion, LW = lecture and discussion with worksheet, SD = self-study and discussion, and SW = self-study and discussion with worksheet. For each LD, LW, SD, and SW group, total, verbatim type, and transfer type items scores are given.

Table 2. Learning outcomes comparison for offline and online groups.

Type of items	Offline		Online		
	LD ($n = 50$)	SD ($n = 58$)	LD ($n = 40$)	SD ($n = 42$)	P
Total Score (25 points)	12.78 (3.14)	16.57 (3.70)	10.08 (4.67)	14.31 (4.03)	<.001
Verbatim (10 points)	7.36 (1.50)	8.57 (1.40)	6.48 (2.69)	7.67 (1.82)	<.001
Transfer (15 points)	5.42 (2.29)	8.00 (3.12)	3.60 (3.05)	6.64 (3.02)	<.001

Note. Data are shown as mean values (standard deviations). LD = lecture and discussion, SD = self-study and discussion. For each LD and SD group, total, verbatim type, and transfer type items scores are given.
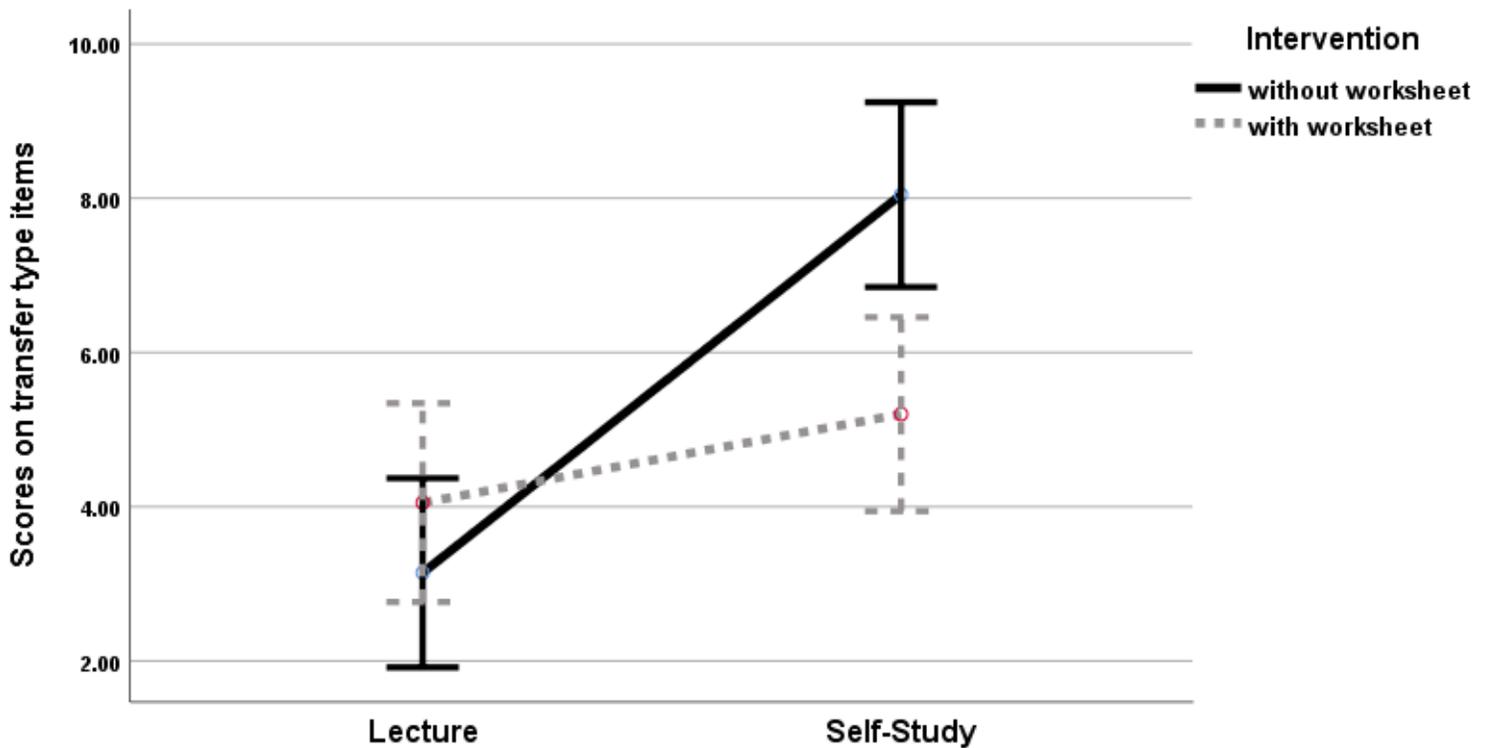
Figures

Figure 1

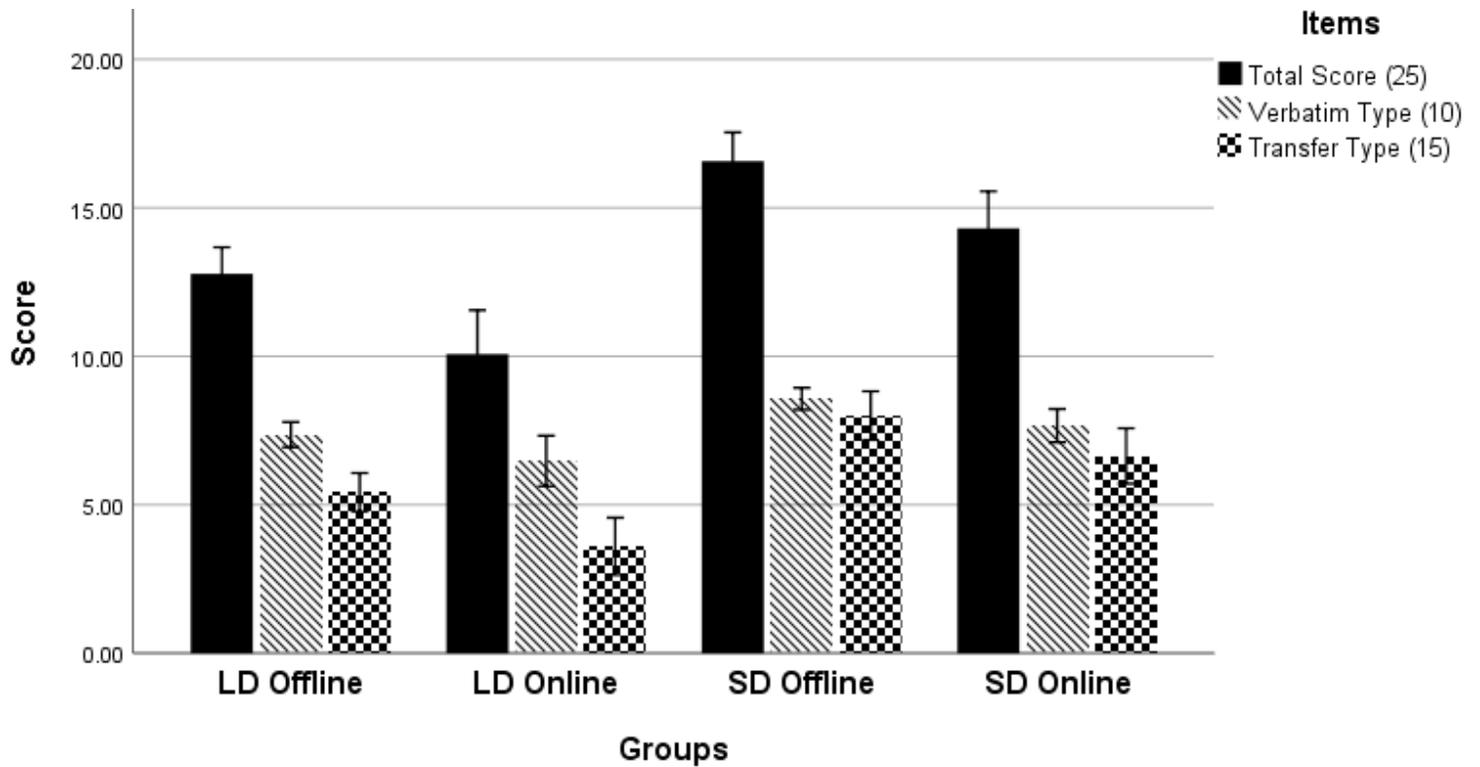The main effects and the interaction effect on transfer type items. Note. Error bars are 95% confidence interval.

Figure 2

Comparisons of the final scores of past (Offline) and present (Online) study. Note. LD = lecture and discussions group; SD = self-study and discussions group. All types of items consist of verbatim type and transfer type items. Error bars are ± 2SE.