Received: 29 May 2020 | Revised: 30 September 2020 | Accepted: 1 November 2020
DOI: 10.1111/jcal.12511

ARTICLE

Journal of Computer Assisted Learning WILEY

Robotic question support system to reduce hesitation for face-to-face questions in lectures

Jiro Shimaya[1] | Yuichiro Yoshikawa[1] | Kohei Ogawa[2] | Hiroshi Ishiguro[1]

[1]Graduate School of Engineering Science, Osaka University, Osaka, Japan

[2]Graduate School of Engineering, Nagoya University, Nagoya, Japan

Correspondence
Jiro Shimaya, Graduate School of Engineering Science, Osaka University, Osaka, Japan.
Email: shimaya.jiro@irl.sys.es.osaka-u.ac.jp

Funding information
Japan Society for the Promotion of Science (JSPS), Grant/Award Numbers: 18H04114, 20H00101; JST-Mirai Program, Grant/Award Number: JPMJMI18C6

Abstract

Encouraging students to actively ask questions during lectures is a formidable challenge that can be addressed through innovative use of information technology. We developed a robotic system that allows students in a lecture to collaboratively decide questions to be asked by a humanoid robot. To verify whether the system reduces hesitation to ask questions during lectures, 62 university students were divided into two groups, and each attended two different mock lectures on the Nobel Prize in Physics. Two lectures were conducted with and without the proposed system in counterbalanced order. Results suggested that students who were usually hesitant to ask questions during lectures became less hesitant to ask questions face-to-face when they could use the proposed system. Moreover, the perceived activeness in the lectures increased when using the system. Multiple regression analyses revealed that certain student actions, particularly tweeting and showing agreement with the questions posted by others, were correlated with an increase in perceived activeness.

KEYWORDS

audience response system, education technology, human–robot interaction, lecture participation support, question support system

1 | INTRODUCTION

In the lectures attended by large number of students, encouraging students to actively ask questions is considered a formidable challenge. Asking questions in class has an important role in enhancing complex knowledge construction (King, 1994) as well as the ability to think critically (Gray, 1993). Psychological factors such as shaming can create hurdles for students who are reluctant to participate in classes (Crozier, 2005; Doyon, 2000; Gong & Yanchar, 2019; Lund, 2008; Qashoa, 2006; Truong, 2017). The perceived in-class passivity of students is regarded as a problem in schools in Japan (Harumi, 2011; Kim, 2019; Leung, 2020) and several other countries (King, 2011; Liu & Littlewood, 1997; Gong & Yanchar, 2019; Truong, 2017). The possibilities of using information and robotic/virtual agent technologies to address such passivity problems have attracted the attention of researchers (Kubilinskiene, Zilinskiene, Dagiene, & Sinkevièius, 2017; Neumann, 2020); for example, in trials, robots or virtual agents have been used as teaching assistants (Alemi, Meghdari, & Ghazisaedy, 2014, 2015; Hameed, Strazdins, Hatlemark, Jakobsen, & Damdam, 2018; Hong, Huang, Hsu, & Shen, 2016; Kanda, Shimada, & Koizumi, 2012), tutors (Alaimi, Law, Pantasdo, Oudeyer, & Sauzeon, 2020; Han, Jo, Jones, & Jo, 2008; Hashimoto, Kobayashi, Polishuk, & Verner, 2013; Saerbeck, Schut, Bartneck, & Janse, 2010), learning partners (Iio et al., 2019; Kory & Breazeal, 2014; Mazzoni & Benvenuti, 2015; Wang, Young, & Jang, 2013) and a schoolmate taught by students (Tanaka & Matsuzoe, 2012).

1.1 | Technology for supporting student questions

Reduced social cues in computer-mediated communication are expected to decrease the pressure in the interaction and increase the amount of disclosure (Joinson, 2001; Sproull & Kiesler, 1986), which would be useful to promote easy and active communication during lectures. Computer-media such as e-mail (Gonglewski, Meloni, & Brant, 2001), chat channels (Tuhkala & Kärkkäinen, 2018) and Twitter

(Tang & Hew, 2017) have been applied to enhance student–student or student–lecturer communication. Despite such positive evaluations for using computer-media in the classroom, previous studies suggested that lectures involving interactions only in the online environment decrease the learning satisfaction of students compared to traditional face-to-face lectures (Kemp and Grieve, 2014; Ebner & Gegenfurtner, 2019). Thus, a combination of traditional face-to-face lecture and computer-mediated communication has been the focus of attention (Yen, Lo, Lee, & Enriquez, 2018).

It has been reported that audience response systems such as Clicker promote the active participation of students by providing another means of asking questions or presenting ideas (Aljaloud, Gromik, Kwan, & Billingsley, 2019; Egelandsdal, Ludvigsen, & Ness, 2019; Lantz & Stawiski, 2014; Mayer et al., 2009). A previous study suggested that the anonymity provided by Clicker enhanced student participation during lectures (Freeman, Blayney, & Ginns, 2006). However, there is a trade-off in using audience response systems: their simple interfaces limit the degree of freedom available to students for presenting questions or ideas, and reduces the flexibility of teacher–student communication. For example, Clicker, a device for selecting given options via clicking, only allows students to submit short, simple responses to the teacher. Although there exist other audience response systems with more complex input methods such as typing or flicking (Damanik, 2020; Nitza & Roman, 2016; Scornavacca, Huff, & Marshall, 2009), they have mainly been used for communication in the question-and-answer format, and still cannot flexibly construct the sequences needed for teacher–student communication. Further, it is sometimes difficult to prepare lecture content that enables students to participate in the lectures using an audience response system (Beatty, Gerace, Leonard, & Dufresne, 2006). Therefore, it is worth developing a novel system that enables students to freely compose their questions and also initiates seamless teacher–student communication that can dynamically cover a wider range of topics that interests them both.

1.2 | Impact of operating robots/virtual agents and its educational applications

Considering the following findings in the studies on virtual experiences (Kumazaki et al., 2019; Nishio, Taura, Sumioka, & Ishiguro, 2013; Rosenberg, Baughman, & Bailenson, 2013), encouraging active participation with response system is conjectured to provide merits not only in the current attendance but also in future one without it. Virtual experiences that simulate flying like a superhero virtual reality technology have been shown to promote prosocial behaviour in users (Rosenberg et al., 2013). When a user controls a humanoid robot via a teleoperation system, and the robot occasionally exhibits facial expressions without being prompted by the user, the feelings of the user are influenced to reflect the given facial expressions (Nishio et al., 2013). It was reported that adolescents with Autism Spectrum Disorders realized the importance of using gestures in communication after engaging in virtual communication experiences with an android robot, whose voice and gestures were controlled to mimic a role of an interviewer (Kumazaki et al., 2019). These studies imply that avatar or robotic technology can support enhanced communication experiences that change user behaviour to reflect a given behaviour. Therefore, by providing students a robotic system that allows them to freely present questions and ideas in class while sensing that they have contributed to the learning experience, the students are expected to change their behaviour to reflect the enhanced experience of active participation with the system; in other words, their hesitation to present questions and ideas in class is expected to decrease.

In this study, we examine the advantage of a robotic question support system, representing a brand new audience response system with which students can collaboratively present questions or ideas in class. With this system, students post candidate questions and ideas to be presented by a humanoid robot placed in the class. When a candidate receives a high number of votes or high agreement from other students, it is presented by the robot. Through such a democratic process, it is expected that the sense of contribution to the presentation of questions and ideas can be felt not only by the student who posted it but also by those who voted for it. A previous field study has shown that the robotic question support system can promote the presentation of questions and ideas via the system (Palinko et al., 2018; Shimaya et al., 2020). However, it is still unclear whether and how direct presentations of questions and ideas by students can be encouraged.

1.3 | Purpose of this study

In this paper, we report the experimental results of mock lectures designed to investigate the effects of the system: specifically, we investigate whether contributing to the presentation of questions and ideas via the system reduces hesitation to the direct presentation of them without the system and promotes a sense of actively participating in the lecture. To enhance the sense of contribution, visual and haptic feedback is communicated to the interface device in order to inform students that their questions or ideas will be presented in the class. Hereafter, questions presented directly by students are called direct questions, while questions presented via the system are called via-system questions.

1.4 | Hypothesis

We hypothesized that the hesitation to ask direct questions would be lower in the lecture with the system than the one without it. Hesitation was evaluated according to the number of questions asked, as well as subjective reports collected immediately after each lecture.

2 | METHOD

2.1 | Participants

Sixty-two first- and second-year Japanese university students were recruited and divided into two groups. Twenty-nine of these students were assigned to Group 1 and first attended an experimental lecture with the proposed system (hereafter referred to as the with-system lecture) and then attended another without it (hereafter referred to as the without-system lecture). The remaining 33 students were assigned to Group 2 and first attended the without-system lecture followed by the with-system lecture. Group 1 consisted of 14 males, 14 females and 1 student who did not disclose their gender, while Group 2 consisted of 17 males, 14 females and 2 students who did not disclose their genders.

Nine less-hesitant participants, whose "hesitation to question in daily lectures" ratings were lower than the neutral level, were excluded from the subjective measure analysis because they were considered to have less need for the system. Namely, it was rated in one of the questionnaire items introduced in Section 2.4, that is, "You hesitate to ask questions during lectures in your daily life," and the "neutral level" meant the intermediate score, that is, 4 ("undecided"), in the questionnaire answer. One participant who answered the questionnaire only partially was also excluded. The remaining 52 participants (26 in Group 1 and 26 in Group 2) were analysed. The average ages and *SD*s of the ages of the participants were 19.5 and 0.8 (min: 18, max: 21) in Group 1 and 19.4 and 0.9 (min: 18, max: 21) in Group 2.

2.2 | System

The robotic question support system consists of an interface device for posting candidate sentences to be spoken and for voting on posted sentences, and a humanoid robot that speaks the elected sentences. The interface device is used to run a Web application and can be accessed from smartphones or personal computers via the target URL. Figure 1 illustrates the appearance of the interface.

To post a sentence to be spoken, a student inputs text and pushes the blue button labelled "question button." To post sentences intended to be shared among students, the yellow button labelled "tweeting button" is pushed instead. The posted sentences are displayed in the interface in the order in which they were posted. The upper section of each post displays text information such as "<tweet>by you @ 2020-03-23 19:13:54," which indicates the type of post (question or tweet), the proposer of the post (this can also be anonymous) and the timestamp of the post.

FIGURE 1 Interface of the robotic question support system [Colour figure can be viewed at wileyonlinelibrary.com]

The bottom section of each post contains a "like" button to allow the student to show agreement with the posted sentence. The number next to the like button represents the number of students who agreed with the sentence. For sentences liked by the student, the number is displayed in red; otherwise, it is displayed in grey. The student can cancel his or her action of liking the sentence by selecting the like button again.

In this experiment, the robot could speak only when the lecturer approved it. The lecturer indicated approval by saying "any questions?" For the simplicity and stability of the experiment, the timing of the approval was recognized by an experimenter. During the approved period, two sentences were randomly chosen from the top-five "liked" sentences and spoken by the robot. Although "liked" tweets can be candidates to be spoken when the number of liked questions is less than two, tweets were not spoken in this experiment.

Before the robot spoke a sentence, the students who contributed to the content were informed that the sentence they posted or liked was going to be spoken. The interface of the student who posted the sentence displayed an alert message such as "Your posted question 'Which physicist do you like the best?' was selected because 15 other students liked it," and the vibrate function on their device was activated three times. The interface of a student who liked the spoken sentence displayed an alert message such as "Your liked question 'Which physicist do you like the best?' was selected with 14 others also liked it," and the vibrate function was activated.

A desktop-type humanoid robot called CommU (Vstone, Co. Ltd.) (Figure 2) was used as the robot for the system. CommU is 30 cm tall and has 14 degrees of freedom of movement to express natural, human-like movement in the upper body during conversations (two in the waist, two in the right shoulder, two in the left shoulder, three in the neck, one in the mouth, three in the eyes and one in the eyelids). The robot speaks through a speaker in its chest while moving its lips synchronously. A commercial software product, AI Talk (AI Inc.), was used for Japanese text-to-speech synthesis. A voice model named *yuuto*, whose voice resembles that of a young boy, was selected in order to match the childlike appearance of the robot.

During lectures, when the robot received no sentence to be spoken, it randomly looked at the lecturer, the blackboard, the students seated on the left and those seated on the right. The targets being looked at were switched every 6–10 s, and the probabilities of the robot looking at the lecturer, the blackboard, the students seated on the left, or those seated on the right were 20, 60, 10 and 10%, respectively. Note that the positions of the targets were predefined. When the robot received a sentence to be spoken, it looked at the lecturer, and then raised its right hand and spoke the sentence.

To encourage the audience to ask direct questions, in this experiment the robot started to speak a sentence 10 s after the lecturer said "Any questions?" If an audience member asked a direct question within 10 s, the robot waited for the next opportunity.

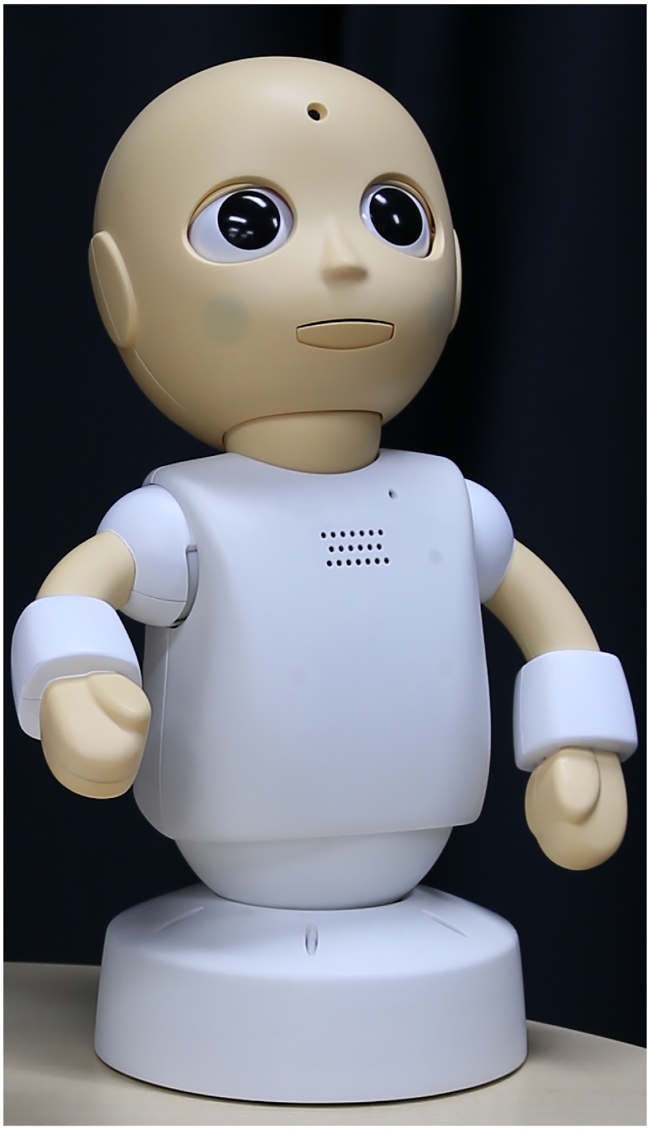

FIGURE 2 A desktop-sized robot CommU [Colour figure can be viewed at wileyonlinelibrary.com]

2.3 | Procedure

In this study, ethical approval was received from the Ethical Committee of the Graduate School of Engineering Science, Osaka University.

Written informed consent was obtained from all participants prior to the experiment. The experiment was conducted in a standard lecture room at the Graduate School of Engineering Science, Osaka University (see Figure 3). Video cameras were placed in the front and back of the room to capture the behaviour of the participants and content from the lecturer, respectively.

The experimental procedure is shown in Figure 4. The experimenter explained the procedure and the rules for behaviour in the experimental lectures. The participants first answered a pre-experiment questionnaire and attended two lectures approximately 40 min in length, including Q&A sessions. Immediately after each lecture, they answered a post-lecture questionnaire. The participants read all questionnaire items before the first lecture session. With-system and without-system lectures were conducted, and their order was counterbalanced. Namely, participants in Group 1 first attended a without-system lecture followed by a with-system lecture, while those in Group 2 first attended the with-system lecture followed by the without-system lecture.

Before the with-system lecture, the participants were provided with ASUS ZenFone Max (M2) smartphones and a CommU robot was placed on a table between the participants' seats and the blackboard. The experimenter explained the usage of the system and allowed participants to attempt to use it. Before the lecture, the experimenter ensured that all participants could log in to the Web page for the interface and post-test questions. For analysis purposes, the user IDs corresponding to the seats were assigned to the participants. The messages sent from or to the smartphones were logged with the user ID, message type (i.e., question, tweet, or like from participants as well as feedback from the system), message text and timestamp of the message.

In both lectures, a lecturer provided an introduction to physics by discussing the Nobel Prize. The lecturer was Mr. Yobinori Takumi, a well-known Japanese lecturer who actively uploads educational material to YouTube to explain such content to general populations of college students. In the first lecture, 10 studies that received Nobel Prizes from 1901 to 1910 were described, while 9 such studies from 1911 to 1920 were presented in the second lecture. In each lecture, Q&A sessions on the previously discussed Nobel Prize studies were inserted approximately every 10 min. The lecturer began the Q&A by saying “do you have any questions?” The Q&A sessions ended after approximately 3 min.

The participants were informed that they would attend two approximately 40-min lectures on the Nobel Prize for the experimental purpose of investigating how the proposed system was used during a lecture. The experimenter explained that they could ask the lecturer questions at any time during the lecture. In addition, for the with-system lecture, the experimenter explained the following rules:

- The participants could ask the lecturer any questions at any time during the lecture, even in the with-system lecture.
- They could post or vote on questions and tweets at any time during the lecture.
- In the Q&A sessions, the robot would randomly choose two questions from the posted questions receiving the many votes, and would use its speech capabilities to ask the questions.
- In the Q&A sessions, if only a few questions had received large numbers of votes, the robot might instead convey tweets receiving high vote totals.
- The posted questions and tweets were erased after the Q&A session.

2.4 | Measurement

2.4.1 | Behavioural measure

As a behavioural measure, instances of the following behaviours were counted:

FIGURE 3 The scene of the lecture in the experiment [Colour figure can be viewed at wileyonlinelibrary.com]

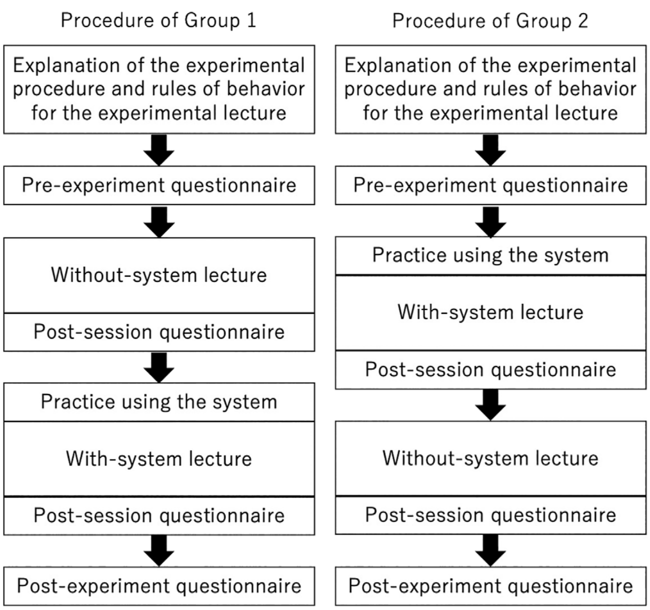

FIGURE 4 Procedure of the experiment

- Number of direct questions.
- Number of questions, tweets and likes submitted via the system.

The direct questions were counted from the video of the lecture, while the via-system questions, tweets and likes were counted from the operation interface log.

2.4.2 | Subjective measure

The questionnaire items used in the current experiment are listed in Table 1. As a subjective measure, "Hesitation," "Activeness," "Interest," "Understanding," "Usefulness of the direct questions [/the via-system questions]," and "Nuisance of the direct questions [/the via-system questions]" were obtained from the post-session questionnaire and the "Daily hesitation" score was obtained from the pre-experiment questionnaire. All items were answered using a seven-point Likert scale (1. I strongly disagree, 2. I disagree, 3. I slightly disagree, 4. Undecided, 5. I slightly agree, 6. I agree, 7. I strongly agree).

In addition to the evaluation by participants, the lecturer also evaluated the following aspects of the lecture after each session: "Lecture quality," "Student activeness" and "Usefulness of the student direct questions [/the student via-system questions]." These items were also answered with the same seven-point Likert scale as described above.

The questionnaire for students was based on the previously used instruments (Shimaya et al., 2020) and the one for the lecturer was developed as part of this study. Other than the first question, the other items were taken from the questionnaire for students for comparison purposes. The first item was to check if the lecturer failed to deliver the lecture as a kind of manipulation check, and therefore such a casual question sentence was chosen.

3 | RESULTS

3.1 | Behavioural measure

The via-system questions were presented twice during every Q&A session in the with-system lectures. No direct questions were asked in the with-system lectures, while one was asked in the without-system lectures. Specifically, one direct question was asked in the without-system lecture for Group 2. Although the participants were allowed to ask direct questions outside of the Q&A session as long as they did so without the system, none did.

3.2 | Subjective measure

The mean, *SD*, range, median and skewness of the answer scores of "Hesitation," "Activeness," "Interest" and "Understanding" are listed in

TABLE 1 Questionnaire items used in the experiment

Timing (who and when)	Label	Sentence
Post-session (students)	Hesitation	You hesitated to ask direct questions during the lecture
	Activeness	You actively participated in the lecture
	Interest	The lecture was interesting
	Understanding	The lecture was understandable
	Usefulness of the direct questions [/the via-system questions]	The direct questions [/the via-system questions] were useful in making the lecture understandable
	Nuisance of the direct questions [/the via-system questions]	The direct questions [/the via-system questions] caused nuisance
Pre-experiment (students)	Daily hesitation	You hesitate to ask questions during lectures in your daily life
Post-session (lecturer)	Lecture quality	You conducted a good lecture
	Student activeness	The students were active during the lecture
	Usefulness of the student direct questions [/the student via-system questions]	The direct questions [/the via-system questions] from the students were helpful in facilitating your lecture

Table 2. These values do not necessarily guarantee that the obtained data are normally distributed. However, studies in recent decades suggest that if the sample sizes are not very small, the parametric tests are robust even if the data are not normally distributed (De Winter, 2013; De Winter & Dodou, 2010; Fagerland, Sandvik, & Mowinckel, 2011; Norman, 2010; Rasch & Guiard, 2004). Thus, we used the *t*-test for comparison in this study. The *t*-test result and Cohen's *d* effect size (Cohen, 1988) are also listed in Table 2. Note that we used the definition of Cohen's *d*, which considers dependency among samples.

The mean score of "Hesitation" was significantly lower in the with-system lectures (M = 5.3, *SD* = 1.5) than in the without-system lectures (M = 5.8, *SD* = 1.2) ($p = 0.038$). The mean score of "Activeness" was significantly higher in the with-system lectures (M = 4.8, *SD* = 1.3) than in the without-system lectures (M = 4.2, *SD* = 1.6) ($p = 0.014$). No significant differences were shown in the mean scores of "Interest" or "Understanding."

The impressions ("Usefulness" and "Nuisance") of the via-system questions were compared with the neutral level 4. The score of "Usefulness of the via-system question" ranged from 1 to 7, and its mean, *SD*, median and skewness were 5.2, 1.5, 5.0 and − 1.10, respectively. The one-sample *t*-test showed that the mean was significantly higher than the neutral level ($t[51] = 6.01$, $p < .001$, Cohen's $d = 2.36$). We used the definition of Cohen's *d* for a one-sample case (Cohen, 1988). The score of "Nuisance of the via-system question" ranged from 1 to 7, and its mean, *SD*, median and skewness were 3.1, 1.5, 3.0 and 0.48, respectively. The one-sample *t*-test showed that the mean was significantly lower than the neutral level ($t[51] = 4.17$, $p < .001$, Cohen's $d = 1.63$). The impressions of the direct questions were not analysed because very few were asked.

The mean scores of the lecturer's self-evaluation on "Lecture quality" were 5.5 and 5 for the with- and without-system lectures, respectively. The mean scores of "Student activeness" as rated by the lecturer were 5 and 2 for the with- and without-system lectures, respectively. The mean score of "Usefulness of the student via-system questions" was 6.5.

3.3 | Regression analysis

For examining how the system influences changes in the mental state of users, multiple regression analyses using the backward-forward stepwise method were applied, using the amount of change in "Hesitation" and "Activeness" as the dependent variables. The independent variables were the number of times system actions such as "question," "tweet," "like for question" and "like for tweet" were logged, and the number of chosen sentences that one posted or liked (i.e., the number of times feedback was received from the system). Table 3 shows the correlation coefficient of each independent variable, standardized partial regression coefficient (β), squared multiple correlation coefficient (R^2), adjusted-R^2 and total *F*-value of the stepwise multiple regression when the dependent variable is the amount of change in "Activeness." The numbers of "tweet" and "like for question" actions are significantly correlated with the amount of change in "Activeness" (adjusted $R^2 = 0.3279$, $p < .001$). The variance inflation factor (VIF) between "tweet" and "like for question" was 1.07, which was less than 10, implying no multicollinearity. The standardized partial regression coefficient of "tweet" and "like for question" are both significant at 0.169 and 0.168, respectively.

TABLE 2 Comparison between the questionnaire scores from the with- and without-system lectures: mean, *SD*, range (min, max), median and skewness of the scores for each lecture type, and the results of the paired *t* test: *t* value (*t*), degree of freedom (df), *p*-value and Cohen's *d* effect size among samples

		Condition		Paired *t* test		
		With-system	Without-system	*t* (df = 51)	*p*	Cohen's *d*
Hesitation	Mean	5.3	5.8	2.12	0.038	0.30
	SD	1.5	1.2			
	Range	(1, 7)	(2, 7)			
	Median	5	6			
	Skewness	−0.66	−1.00			
Activeness	Mean	4.8	4.2	2.54	0.014	0.35
	SD	1.3	1.6			
	Range	(1, 7)	(1, 7)			
	Median	5	4			
	Skewness	−0.75	−0.14			
Interest	Mean	6.0	6.0	0	1.000	0.00
	SD	1.2	0.9			
	Range	(1, 7)	(3, 7)			
	Median	6.0	6.0			
	Skewness	−1.98	−0.78			
Understanding	Mean	5.7	5.8	0.40	0.690	0.06
	SD	1.2	1.0			
	Range	(1, 7)	(3, 7)			
	Median	6.0	6.0			
	Skewness	−1.48	−0.68			

TABLE 3 Correlation coefficient of each independent variable, standardized partial regression coefficient (β), squared multiple correlation coefficient (R^2), adjusted-R^2 and total *F*-value of the stepwise multiple regression with "the amount of change in Activeness" as the dependent variables

			Results of multiple regression analyses with the stepwise method		
Independent variable	Correlation coefficient	β	R^2	Adjusted-R^2	Total *F*-value (df = [2,49])
Tweet	.483	0.169**	0.354	0.328	13.44***
Like for question	.463	0.168**			
Question	.403	N/A			
Like for tweet	.401	N/A			
Feedback	.276	N/A			

Abbreviation: df, degree of freedom; N/A, not available.
$p < .01$.*$p < .001$.

4 | DISCUSSION

4.1 | Decreased hesitation in asking direct question

Previous studies showed decreased hesitation in asking via-system questions (Shimaya et al., 2020) when using question-support systems such as Clicker (Aljaloud et al., 2019; Egelandsdal et al., 2019; Lantz & Stawiski, 2014; Mayer et al., 2009) or other more complex input methods (Damanik, 2020; Nitza & Roman, 2016; Scornavacca et al., 2009). Similarly, Twitter (Tang & Hew, 2017) or chat channels (Tuhkala & Kärkkäinen, 2018) were also found to promote online student–lecturer interaction. Unlike these previous works, the current experiment also shows decreased hesitation in asking direct questions and increased activeness during the with-system lectures. Thus, the hypothesis that "hesitation in asking direct questions will be less in lectures with the system than in those without it" was supported in the subjective measure.

A previous study has shown that robotic/virtual agents could promote students' asking questions by giving hint words for questioning through dialogue (Alaimi et al., 2020). Rather than directly helping the students with the questions to be asked, the current study shows that hesitation in asking questions can be reduced by indirectly giving collaborative opportunities to participate in asking questions to a human teacher via a robot. The result is regarded to imply that students in the with-system lecture felt that they had actively participated in the lecture and could have had self-confidence to do so even without the system. The latter is consistent with previous findings stating that experiences enhanced with a robot change the user's mental state to reflect the behaviour of the robot (Kumazaki et al., 2019; Nishio et al., 2013; Rosenberg et al., 2013). This experiment suggests that such a change could occur not only when users operate a robot alone but also when they do so as part of a group.

On the other hand, because few direct questions were asked regardless of lecture type, it is unclear whether the system would contribute to actual behaviour change. The previous field study that tested a prototype system during lectures on a different topic, namely the introduction of recent advances in robotics, obtained approximately 10 direct questions (Palinko et al., 2018). Perhaps it was easier for students to relate a lecture on the recent advances in robotics to everyday life than one on the history of physics. It is also possible that, during the experimental lecture, the participants felt less necessity and curiosity, which are supposed to be felt to a greater degree in lectures in their daily school lives or those on more casual topics; thus, the motivation to ask questions may have been lower than in the previous field experiment. Extending the system to support students in relating lecture content to their potential interests is a possible and important future study topic.

Studies on digital interactions have coped with the “poor get richer effect”; for example, introverted persons can easily express themselves in online environments compared to traditional face-to-face environments (Amichai-Hamburger, 2007; Amichai-Hamburger & Hayat, 2005; Blau, Weiser, & Eshet-Alkalai, 2017). The proposed messaging technique offers anonymity for hesitant participants demonstrating the “poor get richer effect.” In contrast, there is a concern that it might be an inefficient means of communication for extroverted persons. Although the current system conveyed the message in anonymous ways, some extroverted students might benefit from more anonymous ways. Therefore, it will be interesting to examine the relationships among psychological backgrounds such as introvert/extrovert of participants, the anonymity level of the system, and its benefits as future work.

4.2 | Participants' impression of the lecture

Participant impressions of the via-system questions suggest that these questions did not interfere with the lecture, and were useful in deepening their understanding of the lecture. In addition, the “Interest” and “Understanding” of the with-system lecture were not significantly lower than those of the without-system lecture. On the other hand, the lecturer rated the with-system lecture relatively high in terms of quality and student activeness. He also felt that the with-system questions were helpful in facilitating the lecture. The with-system lecture was positively evaluated by both the lecturer and the students. The results motivate us to move to the next step, which will involve verifying the usefulness of the system in an actual educational setting.

4.3 | Influential options of via-system participation

The results of regression analysis revealed that students who invoked the “tweet” and “like for question” actions more often tended to evaluate themselves as relatively more active in the with-system lecture than in the without-system lecture. Our results suggest that even actions that do not directly result in robot speech, such as “tweet” and “like for questions,” can contribute to enhancing a sense of participation. Identifying the cognitive processes that exist between these actions and the user's mental state is an important step in developing more effective systems.

5 | LIMITATIONS

It is not clear whether the enhanced participation was caused by the presence of the robot or by the use of the interface, or whether it was caused by a combination of these. Previous research showed that when a robot operated by a third person is included in a dialogue between humans, the number of dialogue increases (Kim et al., 2013). This implies that robots have the ability to activate human–human conversation. In this experiment, there were scenes in which the robot's comments prompted the participants to laugh and send tweets expressing familiarity with the robot's behaviour. The robot's behaviour and presence itself may have contributed to the ease of asking face-to-face questions.

Concerns about being negatively evaluated by others and lack of confidence in the content of questions are considered to be the reasons for Japanese university students to hesitate in asking questions during lectures (Akita, 1995). On the other hand, being able to check others' posts and likes in advance through the operation interface can reduce the uncertainty of other participants' responses. Thus, the operation interface itself reduced concerns about negative evaluations and increased the ease of asking questions without the system. Therefore, further experiments are necessary to distinguish between the effect of operating the robot and the effect of communicating via the interface.

These results were obtained from mock lectures on the Nobel Prize in Physics, attended by approximately 30 first- and second-year university students. It is unclear to what degree the participant group size, participant characteristics and lecture theme influenced the effects observed in this experiment. In addition, because all participants were new to the system in this experiment, the effects of using it for longer periods of time were not examined. Conducting long-term experiments involving a more varied field of participants and lecture themes are important for future work.

6 | CONCLUSION

We developed a system that enables multiple lecture attendees to collaboratively select sentences to be spoken by a robot to enhance lecturer–student communication. The results of a mock lecture experiment verified that students who were usually hesitant to ask questions during lectures became less hesitant to ask questions face-to-face when using the system. Moreover, the perceived activeness in lectures was increased with the proposed system. These positive results encourage us to conduct a further longitudinal study to utilize the proposed system to encourage students' active participation in an actual classroom. As the effect on behavioural aspects was unclear in the experimental setting of the study, it is worth extending the system to support students in finding questions to be asked. Beyond the earlier cognitive studies on teleoperation, the current study reports that the operator's attitude can be influenced by the operated agent even when operated by multiple persons. Further, future work should be dedicated to clarifying the cognitive model behind student participation engendered by the system, and testing the system in actual educational environments.

ACKNOWLEDGEMENTS

We would like to thank Mr. Yobinori Takumi for giving the mock lecture and useful feedback for the study. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 18H04114, 20H00101 and JST-Mirai Program Grant Number JPMJMI18C6, Japan.

CONFLICT OF INTEREST

The authors have no conflicts of interest directly relevant to the content of this article.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12511.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Jiro Shimaya https://orcid.org/0000-0003-1794-6567

REFERENCES

Akita, K. (1995). Perspectives on psychology and questioning behavior: A comparison of general education courses and psychology majors. *Annual Report of the Department of Psychology, Rikkyo University*, (38), 25–38 (in Japanese).

Alaimi, M., Law, E., Pantasdo, K. D., Oudeyer, P. Y., & Sauzeon, H. (2020, April). Pedagogical agents for fostering question-asking skills in children. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–13).

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2014). Employing humanoid robots for teaching english language in iranian junior high-schools. *International Journal of Humanoid Robotics*, *11*(3), 1450022.

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2015). The impact of social robotics on l2 learners' anxiety and attitude in english vocabulary acquisition. *International Journal of Social Robotics*, *7*(4), 523–535.

Aljaloud, A. S., Gromik, N., Kwan, P., & Billingsley, W. (2019). Saudi undergraduate students' perceptions of the use of smartphone clicker apps on learning performance. *Australasian Journal of Educational Technology*, *35*(1), 85–99. https://doi.org/10.14742/ajet.3340.

Amichai-Hamburger, Y. (2007). Personality, individual differences and Internet use. In *The Oxford handbook of internet psychology* (pp. 187–204). Oxford, England: Oxford University Press.

Amichai-Hamburger, Y., & Hayat, Z. (2005). Personality and the internet. In *The social net: Human behavior in cyberspace* (pp. 27–55). Oxford, England: Oxford University Press.

Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics*, *74*(1), 31–39.

Blau, I., Weiser, O., & Eshet-Alkalai, Y. (2017). How do medium naturalness and personality traits shape academic achievement and perceived learning? An experimental study of face-to-face and synchronous e-learning. *Research in Learning Technology*, *25*. https://doi.org/10.25304/rlt.v25.1974

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Crozier, W. R. (2005). Shyness and students' perceptions of seminars. *Psychology Learning & Teaching*, *4*(1), 27–34.

Damanik, E. S. D. (2020). Student attitude toward the use of Whatsapp in EFL class. *Vision*, *15*(2), 13–22.

De Winter, J. C. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research, and Evaluation*, *18*(1), 10.

De Winter, J. F. C., & Dodou, D. (2010). Five-point likert items: T test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research, and Evaluation*, *15*(1), 11. https://doi.org/10.7275/bj1p-ts64 (Addendum added October 2012).

Doyon, P. (2000). Shyness in the Japanese EFL class: Why it is a problem, what it is, what causes it, and what to do about it. *The Language Teacher*, *24*(1), 11–16.

Ebner, C., & Gegenfurtner, A. (2019). Learning and satisfaction in webinar, online, and face-to-face instruction: A meta-analysis. *Practical Assessment, Research, and Evaluation*, *15*(1), 11. https://doi.org/10.3389/feduc.2019.00092

Egelandsdal, K., Ludvigsen, K., & Ness, I. J. (2019). Clicker interventions in large lectures in higher education. In *Learning, design, and technology*. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-17727-4_147-1

Fagerland, M. W., Sandvik, L., & Mowinckel, P. (2011). Parametric methods outperformed non-parametric methods in comparisons of discrete numerical variables. *BMC Medical Research Methodology*, *11*(1), 44. https://doi.org/10.1186/1471-2288-11-44

Freeman, M., Blayney, P., & Ginns, P. (2006). Anonymity and in class learning: The case for electronic response systems. *Australasian Journal of Educational Technology*, *22*(4), 568–580. https://doi.org/10.14742/ajet.1286

Gong, S. P., & Yanchar, S. C. (2019). Question asking and the common good: A hermeneutic investigation of student questioning in moral configurations of classroom practice. *Qualitative Research in Education*, *8*(3), 248–275. https://doi.org/10.17583/qre.2019.3947

Gonglewski, M., Meloni, C., & Brant, J. (2001). Using e-mail in foreign language teaching: Rationale and suggestions. *The Internet TESL Journal*, *7*(3), 1–12.

Gray, P. (1993). Engaging students' intellects: The immersion approach to critical thinking in psychology instruction. *Teaching of Psychology*, *20*(2), 68–74.

Hameed, I. A., Strazdins, G., Hatlemark, H. A., Jakobsen, I. S., & Damdam, J. O. (2018, February). Robots that can mix serious with fun. International conference on advanced machine learning technologies and applications (pp. 595–604). Cham: Springer. https://doi.org/10.1007/978-3-319-74690-6_58

Han, J.-H., Jo, M.-H., Jones, V., & Jo, J.-H. (2008). Comparative study on the educational use of home robots for children. *Journal of Information Processing Systems*, *4*(4), 159–168.

Harumi, S. (2011). Classroom silence: Voices from Japanese EFL learners. *ELT Journal*, *65*(3), 260–269.

Hashimoto, T., Kobayashi, H., Polishuk, A., & Verner, I. (2013). Elementary science lesson delivered by robot. 2013 8th ACM/IEEE international conference on human-robot interaction (HRI) (pp. 133–134).

Hong, Z.-W., Huang, Y.-M., Hsu, M., & Shen, W.-W. (2016). Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *Journal of Educational Technology & Society*, *19*(1), 337–349.

Iio, T., Maeda, R., Ogawa, K., Yoshikawa, Y., Ishiguro, H., Suzuki, K., … Hama, M. (2019). Improvement of Japanese adults' English speaking skills via experiences speaking to a robot. *Journal of Computer Assisted Learning*, *35*(2), 228–245.

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, *31*(2), 177–192. https://doi.org/10.1002/ejsp.36

Kanda, T., Shimada, M., & Koizumi, S. (2012). Children learning with a social robot. 2012 7th ACM/IEEE international conference on human-robot interaction (HRI) (pp. 351–358).

Kemp, N., & Grieve, R. (2014). Face-to-face or face-to-screen? Undergraduates' opinions and test performance in classroom vs. online learning. *Frontiers in psychology*, *5*, 1278. https://doi.org/10.3389/fpsyg.2014.01278.

Kim, D. K. (2019). Enhancing Japanese student engagement: The voice of experience. *Bulletin of Institute for Education and Student Services, Okayama University*, *4*, 124–133. http://doi.org/10.18926/58041

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038–1049.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal*, *31*(2), 338–368.

King, J. E. (2011). *Silence in the second language classroom*. (Unpublished doctoral dissertation). University of Nottingham.

Kory, J., & Breazeal, C. (2014). Storytelling with robots: Learning companions for preschool children's language development. The 23rd IEEE international symposium on robot and human interactive communication (pp. 643–648).

Kubilinskiene, S., Zilinskiene, I., Dagiene, V., & Sinkevièius, V. (2017). Applying robotics in school education: A systematic review. *Baltic Journal of Modern Computing*, *5*(1), 50–69. https://doi.org/10.22364/bjmc.2017.5.1

Kumazaki, H., Muramatsu, T., Yoshikawa, Y., Matsumoto, Y., Ishiguro, H., Mimura, M., & Kikuchi, M. (2019). Role-play-based guidance for job interviews using an android robot for individuals with autism spectrum disorders. *Frontiers in Psychiatry*, *10*, 239.

Lantz, M. E., & Stawiski, A. (2014). Effectiveness of clickers: Effect of feedback and the timing of questions on learning. *Computers in Human Behavior*, *31*, 280–286.

Leung, R. C. Y. (2020). Understanding the in-class behavior patterns of Japanese university students in EFL classes: A foreign English teacher's perspective. *Jissen Women's University CLEIP Journal*, , (6), 9–22. https://doi.org/10.34388/1157.00002129

Liu, N.-F., & Littlewood, W. (1997). Why do many students appear reluctant to participate in classroom learning discourse? *System*, *25*(3), 371–384.

Lund, I. (2008). 'I just sit there': Shyness as an emotional and behavioural problem in school. *Journal of Research in Special Educational Needs*, *8*(2), 78–87.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., … Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, *34*(1), 51–57.

Mazzoni, E., & Benvenuti, M. (2015). A robot-partner for preschool children learning english using socio-cognitive conflict. *Journal of Educational Technology & Society*, *18*(4), 474–485.

Neumann, M. M. (2020). Social robots and young children's early language and literacy learning. *Early Childhood Education Journal*, *48*(2), 157–170. https://doi.org/10.1007/s10643-019-00997-7

Nishio, S., Taura, K., Sumioka, H., & Ishiguro, H. (2013). Teleoperated android robot as emotion regulation media. *International Journal of Social Robotics*, *5*(4), 563–573.

Nitza, D., & Roman, Y. (2016). WhatsApp messaging: Achievements and success in academia. *International Journal of Higher Education*, *5*(4), 255–261. https://doi.org/10.5430/ijhe.v5n4p255

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*((5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Palinko, O., Shimaya, J., Jinnai, N., Ogawa, K., Yoshikawa, Y., & Ishiguro, H. (2018). Improving teacher-student communication during lectures using a robot and an online messaging/voting system. 2018 27th IEEE international symposium on robot and human interactive communication (pp. 1–7). https://paperpile.com/app/p/a8c5f61f-20de-0951-a1e3-419b87ce9833

Qashoa, S. H. H. (2006). Motivation among learners of English in the secondary schools in the eastern coast of the UAE.

Rasch, D., & Guiard, V. (2004). The robustness of parametric statistical methods. *Psychology Science*, *46*, 175–208. https://doi.org/10.1214/aoms/1177728845

Rosenberg, R. S., Baughman, S. L., & Bailenson, J. N. (2013, January). Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PLoS One*, *8*(1), e55003.

Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor. Proceedings of the Sigchi conference on human factors in computing systems (pp. 1613–1622).

Scornavacca, E., Huff, S., & Marshall, S. (2009). Mobile phones in the classroom: If you can't beat them, join them. *Communications of the ACM*, *52*(4), 142–146.

Shimaya, J., Yoshikawa, Y., Palinko, O., Ogawa, K., Jinnai, N., & Ishiguro, H. (2020). Active participation in lectures via a collaboratively controlled robot. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-020-00651-y

Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science*, *32*(11), 1492–1512. https://doi.org/10.1287/mnsc.32.11.1492

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, *1*(1), 78–95.

Tang, Y., & Hew, K. F. (2017). Using Twitter for education: Beneficial or simply a waste of time? *Computers & Education*, *106*, 97–118. https://doi.org/10.1016/j.compedu.2016.12.004

Truong, N. T. N. (2017). Understanding first year university students' passivity via their attitudes and language behaviors towards answering questions in class. *Journal of Science HCMC OU-Social*

Sciences, *7*(1), 84–93. https://doi.org/10.46223/HCMCOUJS.soci.en.7.1.295.2017

Tuhkala, A., & Kärkkäinen, T. (2018). Using slack for computer-mediated communication to support higher education students' peer interactions during Master's thesis seminar. *Education and Information Technologies*, *23*(6), 2379–2397. https://doi.org/10.1007/s10639-018-9722-6

Wang, Y. H., Young, S. S.-C., & Jang, J.-S. R. (2013). Using tangible companions for enhancing learning english conversation. *Journal of Educational Technology & Society*, *16*(2), 296–309.

Yen, S. C., Lo, Y., Lee, A., & Enriquez, J. (2018). Learning online, offline, and in-between: Comparing student academic outcomes and course satisfaction in face-to-face, online, and blended teaching modalities. *Education and Information Technologies*, *23*(5), 2141–2153. https://doi.org/10.1007/s10639-018-9707-5

How to cite this article: Shimaya J, Yoshikawa Y, Ogawa K, Ishiguro H. Robotic question support system to reduce hesitation for face-to-face questions in lectures. *J Comput Assist Learn*. 2020;1–11. https://doi.org/10.1111/jcal.12511