Peer Reviewed Article

OPEN JOURNALS IN EDUCATION

Volume 9, Issue 2 (2024), pp. 1-20
International Journal of
Multidisciplinary Perspectives in Higher Education
ISSN: 2474-2546 Print/ ISSN: 2474-2554 Online
https://ojed.org/jimphe

Acceptability, Feasibility, and Effectiveness of an Artificial Intelligence Chatbot in an Asynchronous Epidemiology Course

Eric J. Conrad
Austin White
Job M. Ramirez
Alyssa Alvara
Laura K. Rees

California State University, Stanislaus, California, United States

ABSTRACT

This study assessed the implementation of a virtual assistant chatbot, EpiBot, in an online epidemiology course using the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework to evaluate acceptability, feasibility, and effectiveness. EpiBot was piloted by 29 public health undergraduate students at a Hispanic Serving Institution on the West Coast of the United States. EpiBot's use and accuracy were monitored via the Amazon Web Services platform and an electronic questionnaire, based on the Technology Acceptance Model, was used to evaluate participant perceptions of usefulness, ease of use, attitudes, intentions, self-efficacy, and preferences. Evaluation of EpiBot's sustainability in the course was guided by the Program Sustainability Assessment Tool. EpiBot accurately answered 93.7% of 1607 student submitted queries (89.5% epidemiology content; 10.5% course administration). Participants perceived EpiBot as useful and easy to use, reported a positive impact on their grades (65.5%), used EpiBot as their primary source of information (55.2%), and experienced minimal barriers to use. Key implementation and maintenance challenges included time for development and non-developer training; platform limitations; and operational costs. Strengths supporting sustainability included capacity for adaptation, supportive climate, and interdisciplinary collaboration.

Keywords: Chatbot-Mediated Learning, Public Health, Distance Learning, Asynchronous, Artificial Intelligence

Distance learning leverages digital technology to promote learning experiences and engagement with educational or course material without the geographical or temporal limitations as required in a traditional classroom setting (Moore et al., 2011). These traditional classroom limitations are removed through the implementation of communication and digital technologies that facilitate synchronous (real-time) and asynchronous (self-paced) education (Anderson, 2008; Hrastinski, 2008). Prior to 2020, online education saw significant growth due to factors that include advancements in digital technology, cost effectiveness, and increased flexibility and accessibility offered to learners (Bell & Federman, 2013). The prevalence of online learning further increased due to the COVID-19 pandemic and the need for increased virtual learning opportunities. According to the National Center for Education Statistics (2023), in 2018 only 34.7% of college students reported taking at least one online course and 13.3% reported enrollment in fully online programs. Following the onset of the COVID-19 pandemic, 75% of all undergraduate students reported taking at least one online course and 44% reported enrollment in fully online programs (Cameron et al., 2021).

Online education presents numerous benefits that can enhance the educational experience. One significant advantage is expanded educational opportunities, through increased flexibility and accessibility, which allows learners to access materials and content by reducing geographical and time-related limitations. Previous research has also indicated that well-designed online education courses can result in learning outcomes that are comparable or superior to traditional instruction (Cavanaugh & Jacquemin, 2015; Means et al., 2013) and access to online education is perceived as valuable by both students and faculty (Martin et al., 2019). Importantly, online education can also provide opportunities to increase digital literacy skills as well as provide global perspectives and materials which can facilitate a more inclusive and multicultural learning environment (Bawa, 2016).

There are also challenges and limitations of online learning in the college environment which can result in reduced academic performance and overall learning. Students in online courses may experience a sense of isolation due to reduced opportunities for social learning and a lack of face-to-face interaction can impede perceived access to the instructor (Zhang et al., 2020). Success in online courses may also require students to have a high degree of motivation, self-discipline, and time management skills (Broadbent & Poon, 2015). Additionally, technical issues and low levels of digital literacy can also present a significant challenge to online learners (Sun & Chen, 2016). Learners may also have issues with poor course design or a lack of intuitive navigation in online learning platforms. This may lead to difficulties in locating critical course materials and information, which can negatively impact students' learning experiences, motivation, and engagement in the course (Alqurashi, 2019; Ritzhaupt et al., 2014). Another important limitation is a potential lack of immediate access to course instructors which may promote decreased engagement and hesitance to ask questions (Hung et al., 2010). These delays in communication are further amplified in asynchronous courses and can disrupt the learning process by increasing student hesitance to seek clarification and assistance (Martin et al., 2018). This hesitance may be magnified among first-generation college students and students from traditionally marginalized communities due to heightened levels of anxiety, perceived incompetence, perceived lack of instructor empathy, unconscious bias, or other socio-cultural factors (Ma & Shea, 2021). Such hesitancy in seeking help may impede academic performance and overall learning experience.

One potential solution to facilitate online instruction and student learning is the development and implementation of virtual assistants or chatbots that simulate conversation with text or voice inquiries to generate meaningful answers. Chatbots handle a wide array of inquiries without human input and can synchronously respond to both general and personalized inquiries in real-time, 24 hours per day, with the capacity to be immensely scalable and adaptable (Winkler & Söllner, 2018, p. 23). There has been increased integration of chatbots in the education setting, with reported uses in the domains of teaching and learning, administration, and assessment. Within the domain of teaching and learning, educational chatbots have been reported to provide accurate information through question and response, enhance student support, provide personalized learning, expedite access to materials, and increase learner interest (Crutzen et al., 2011; Hiremath et al., 2018; Mikic-Fonte et al., 2018; Sinha et al., 2020; Song et al., 2017; Verleger & Pembridge, 2018). Additionally, chatbots have been shown to reduce the administrative burden on educators as automation of routine tasks, such as answering FAQs, allows instructors more time to focus on content delivery and individualized instruction (Kuhail et al., 2023; Winkler & Söllner, 2018). Lastly, in a systematic review of chatbot applications in education, Okonkwo and Ade-Ibijola (2021) indicate that chatbots benefit the educational system through the integration of all course or topic-related information, quick access to educational information, motivation and engagement through interactive systems, capacity to handle multiple users or queries simultaneously, and immediate assistance or support. The authors also report challenges related to ethics, evaluation, user attitude, programming, as well as supervision and maintenance.

While previous research has shown beneficial application of chatbots in education (Cunningham-Nelson et al., 2019; Smutny & Schreiberova, 2020; Thomas, 2020; Winkler & Söllner, 2018) further implementation and research is needed. Hwang et al. (2020) indicate the need for application in seldom-applied domains which may include nursing and public health-related fields. Additionally, research that evaluates the implementation, feasibility, acceptability, and attitudes towards use, and adaptability among diverse college populations is needed (Følstad & Brandtzæg, 2017; Kuhail et al., 2023; Wollny et al., 2021). Future research should also seek to employ systematic evaluation methods and provide details regarding development and evaluation of technologies that allow non-programmers to develop and implement chatbots (Kuhail et al., 2023). In order to add to the existing literature, the purpose of this study was to evaluate the acceptability, feasibility, and effectiveness of a pilot implementation of a virtual assistant chatbot (EpiBot) in an asynchronous, online epidemiology course at a Hispanic-Serving Institution (HSI).

Methods

Settings and Participants

The pilot study was conducted at a medium-sized university located on the West Coast of the United States. Participants were recruited from students enrolled in an online, basic epidemiology course during one academic semester. To be eligible to enroll in the course, students needed to be declared department majors, junior or senior academic status, and have completed prerequisite courses such as introductory statistics. Students were directly recruited at the beginning of the semester via campus email and through study-specific advertisements posted in the course Learning Management System (LMS), Blackboard. Study-specific materials included 1) an information sheet regarding the study, 2) link to the website hosting the chatbot called EpiBot and 3) a brief tutorial video demonstrating how to use EpiBot. Using EpiBot for the course was voluntary and students could opt-out or choose to not participate in the study at any time. Students opting to use EpiBot were required to agree to a statement indicating their responses were anonymous but would be used for the purpose of the study. Participants were required to agree to these statements prior to accessing EpiBot initially. Informed consent of participants was obtained for the evaluation questionnaire at the end of the semester. The protocol for this study was approved by the university institutional review board.

Chatbot Development and Application

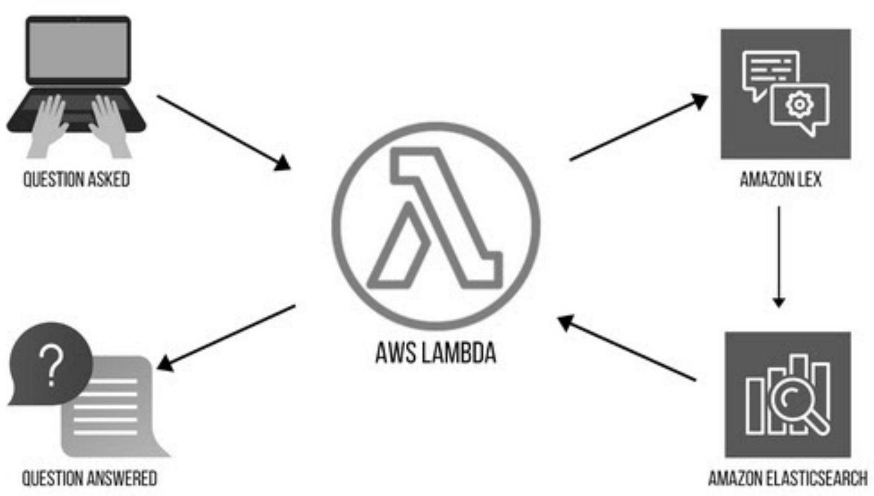
EpiBot, the virtual assistant chatbot, was developed to assist students with course content (epidemiology) and answer questions related to course administration (instructor information, due dates, course policies). The initial iteration of EpiBot was designed to address content corresponding to 10 course modules: (1) Introduction to Epidemiology; (2) History of Epidemiology; (3) Theoretical Foundations of Epidemiology; (4) Measurement of Applied Epidemiology; (5) Surveillance; (6) Outbreak Investigation; (7) Epidemiological Studies; (8) Behavioral Applied Epidemiology; (9) Information Presentation and Communication; (10) Applied Epidemiological Case Studies. The majority of the information provided in the initial iteration of EpiBot was cited from Principles of Epidemiology in Public Health Practice, 3rd edition (Centers for Disease

Control and Prevention, 2011). Course-related information and policies were taken from the course syllabus and input into EpiBot.

EpiBot was developed using the Amazon Web Services (AWS) platform. AWS is a cloud computing service owned by Amazon; companies or individuals can use AWS to host websites, databases, or processes that are needed in an information technology (IT) infrastructure. The initial build of EpiBot utilized a Question and Answer (QnA) template (Strahan, 2017) which leverages Amazon Elasticsearch, AWS Lambda, Amazon Lex, and Amazon S3 for a determined conversational path. Using this template of features, EpiBot was hosted on the AWS platform and able to be accessed via a Uniform Resource Locator (URL) link. Developers had access to a content designer where responses could be created to address user intents and utterances. Amazon Elasticsearch stores the QnA data and matches user intents (questions) to appropriate responses (answers) found in the built data. Amazon Lex is used for natural language processing which allows the chatbot to read, interpret, and understand user requests. AWS Lambda is the connection between the front-end user input and the back-end bot capabilities and Amazon S3 is used to store data files such as the website code for EpiBot. Figure 1 provides a representation of the interaction between these components. When the user inputs an intent or question in the EpiBot front-end, AWS Lambda transfers this data to Amazon Lex for interpretation. Once Lex deciphers the intent, the data is transferred to Amazon Elasticsearch to select the appropriate response for the intent. Responses are then returned to the EpiBot front-end interface through AWS Lambda where users can use the information to supplement their course materials and interact with provided URLs, resources, and videos. This process is repeated when EpiBot is provided with an additional query.

Figure 1

Representation of EpiBot integration using AWS Lambda, Amazon Lex, and Amazon Elasticsearch.

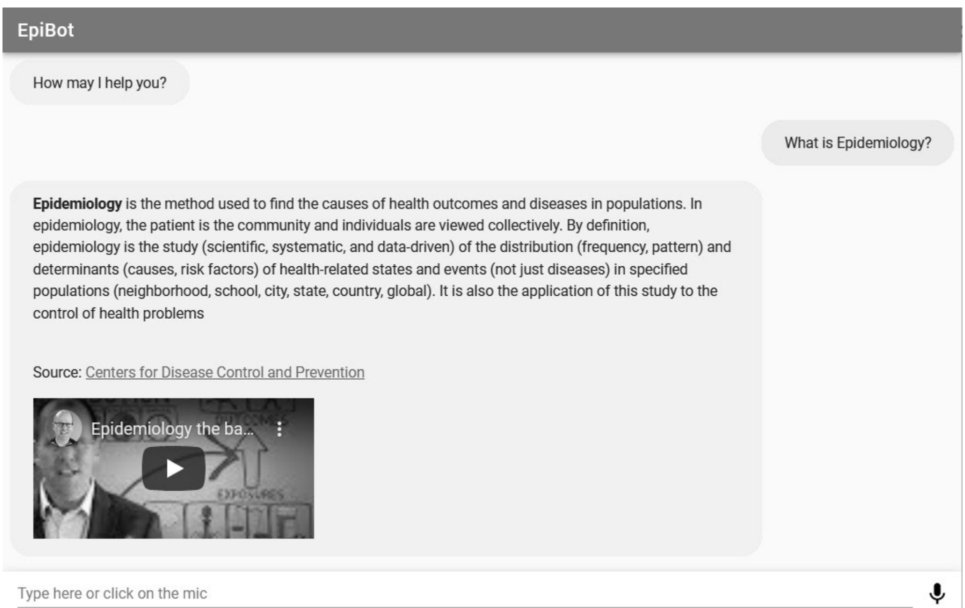


Note. Depiction of EpiBot's workflow: User queries are processed by AWS Lambda, interpreted by Amazon Lex, and matched with an appropriate response from the instructor created repository using Amazon Elasticsearch. The response is routed back through AWS Lambda to the user interface (computer/mobile device).

An example of the front-end user experience for EpiBot is illustrated in Figure 2. Using AWS services, EpiBot allows the user to input text intents or to use voice recognition to speak to EpiBot due to the capability of Amazon Lex to process, read, and play back speech-to-text. Once verbal requests are processed, they are transferred using the same method as a text-based intent. Additionally, the platforms allowed for dynamic functionality in the responses sent to answer user questions. In addition to the ability to provide text-based answers, text could also be linked to an outside URL allowing users to navigate to additional information. It is also possible to include images (e.g tables or figures) or embed videos directly into the responses which can be viewed by the user within the EpiBot interface. Following the completion of the initial iteration of EpiBot, users were able to access the virtual assistant 24 hours a day, seven days a week.

Figure 2

Example of the front-end user experience with EpiBot



Note. Display of the EpiBot user interface on a computer screen. Through this interface, users can submit queries and receive corresponding answers from the repository, which may contain information, definitions, and examples, as well as embedded images, URL links, and videos.

Evaluation Framework

The Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM) implementation science framework was used to facilitate the evaluation of the acceptability, feasibility, and effectiveness of the pilot implementation of EpiBot (Glasgow et al., 1999). Evaluation criteria for each RE-AIM category can be found in Table 1.

Reach was estimated based on the proportion and characteristics of students who indicated agreement to participate in the pilot which was assessed through a questionnaire sent to students at the end of the semester.

Effectiveness was assessed using a theory-based, 49-item (question) electronic questionnaire which employed previously validated scales from the Technology Acceptance Model (TAM) (Gardner & Amoroso, 2004; Park, 2009). TAM scales were modified to reflect chatbot-mediated learning. Modified scales used to assess effectiveness included seven items for perceived ease of use (TAM-PE; e.g. "I find the chatbot easy to use"), seven items for perceived usefulness (TAM-PU; e.g. "using the chatbot can improve my course performance"), four items for attitudes towards the chatbot (TAM-AT; e.g. "I have fun interacting with the chatbot"), four items for behavioral intention (TAM-BI; e.g. "I always try to use the chatbot in as many cases/occasions as possible"), and self-efficacy (TAM-SE; two items). All scale items were rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). One multiple-choice questionnaire item asked students to indicate how their grade was affected as a result of utilizing EpiBot (worse grade; same grade; improved by less than one letter grade; improved by one full letter grade; and improved by more than two letter grades). Effectiveness was also measured by assessing the proportion of correct responses from EpiBot when answering student intents.

Adoption at the individual level was assessed using a survey item of self-reported individual frequency of chatbot use, where a chatbot use was defined as each time a participants used EpiBot and not each individual question asked. The total number of requests obtained from the back-end AWS data were coded as epidemiology-related or administrative content to assess search parameters and determine the proportion of each type of request. To determine factors that may influence adoption, three items were used to assess perceived voluntariness of chatbot use (TAM-PV; e.g. "Using the chatbot is voluntary as far as work is concerned"), four items to assess past experience (TAM-PaEx; e.g. "I have a great deal of experience using chatbots"), and three items to assess perceived complexity (TAM-PC; e.g. "When I use the chatbot, I find it difficult to integrate the results into my course work"). An open-ended questionnaire item was used to assess perceived barriers to using EpiBot. One item asked participants to indicate the platform that they primarily used to access EpiBot (computer, mobile device, both). One item asked participants which course resource they used first when having a course-related question. Multiple choice options included: reading the syllabus or textbook; asking a classmate/peer; contacting the instructor; asking EpiBot; using a search engine (Google, Yahoo, Bing, etc); or other.

Implementation assessed fidelity of the intervention with consideration for consistency, adaptation, time, and cost. One item asked participants to indicate if they experienced technical difficulties when using EpiBot, with an open-ended follow-up item to describe any issues. Additionally, a Google Sheet was created in order for users to report technical issues or incorrect responses from EpiBot. This process was demonstrated in the tutorial video and through reminders in the student LMS. Adaptation was the increase in intent pool questions from initial implementation to the end of the semester. Lastly, perceived time, cost, and barriers were assessed through a tracking log where the faculty and developers could comment on these elements as the intervention progressed.

Maintenance was assessed to determine the extent to which the program could become institutionalized or sustained within the course over time. At the program level, maintenance was assessed qualitatively using the framework of the Program Sustainability Assessment Tool (PSAT) (Luke et al., 2014). The PSAT was selected due to its utility in evaluating implementation of public health and educational programs (Berzosa et al., 2017; Calhoun et al., 2014; Luke et al., 2014; Stoll et al., 2015). Faculty and developers reviewed the eight PSAT domains: (1) environmental support; (2) funding stability; (3) partnerships; (4) organizational capacity; (5) program adaptation; (6) program evaluation; (7) communication; and (8) strategic planning following the pilot semester. Evaluation was focused on identifying strengths and threats as well as to develop strategies for continued maintenance.

Table 1.

Evaluating EpiBot for Acceptability, Feasibility, and Effectiveness Using RE-AIM

RE-AIM Element	**Definition**	**Evaluation Metric**
Reach	The number and percentage of invited/eligible students who participated and their representativeness	Percentage of completed questionnaires. Percentage eligible/excluded (and characteristics) measured via self-report surveys.

Effectiveness	The reported/observed success of EpiBot of perceived student outcomes and EpiBot accuracy	Modified Technology Acceptance Model (TAM): Perceived Easy of Use Scale (TAM-PE); Perceived Usefulness Scale (TAM-PU), Attitudes Scale(TAM-AT), Behavioral Intention Scale (TAM-BI), and Self-Efficacy Scale (TAM-SE). Perceived impact of EpiBot on course grades; Percentage of questions correctly answered by EpiBot.
Adoption (Individual)	Degree of adoption based on individual frequency and intensity of EpiBot use, search choice/parameters, and student reported benefits/barriers	Back-end AWS data (total requests); Proportion of requests epidemiology or course-related content; Proportion of participants who used EpiBot before other options; Perceived Voluntariness Scale (TAM-PV), Past Experience Scale (TAM-PaEp), Perceived Complexity Scale (TAM-PC); Open-ended questionnaire items for barriers
Implementation	Fidelity of intervention elements including consistency, adaptation, time, and cost	Proportion of students who reported technical difficulties and type; adaptation/increase in intent pool from baseline implementation (AWS data); perceived cost (time and resources); Faculty/Developer perceived barriers to implementation
Maintenance	The extent to which the program could become institutionalized or sustained within the course over time	Program Sustainability Assessment Tool (PSAT).

Notes. AWS = Amazon Web Services

Data Analysis

Data from end of semester questionnaires were analyzed using the IBM Statistical Package for the Social Sciences (SPSS, version 27). Descriptive statistics were calculated to describe sample characteristics and TAM Scales. Scale values were derived from the summation of all items within the individual scale.

Results

Complete demographic information is presented in Table 2. Feasibility, effectiveness, and acceptability results are described below by each dimension of the RE-AIM framework.

Reach

Participant characteristics are presented in Table 2. A total of 29 students were enrolled in the basic epidemiology course at the census date indicating the eligible sample to participate in the pilot. All 29 students provided consent to participate, completed the end of semester questionnaire, and reported using EpiBot at least once during the semester. At baseline, participants had a mean age of 21.7 ± 3.0 years. Most participants were female (69%), of Hispanic/Latinx/Spanish origin (51.7%) or non-Hispanic White (37.9%), first-generation college students (79.3%) with prior online course experience (93.1%), and English as their first language (62.1%).

Table 2

Participant characteristic information (n = 29)

Variables	
Age; mean (SD)	21.7(3.0)
Gender; n (%)	
Male	8(27.6)
Female	20(69.0)
Other	1(3.4)
Race/ethnicity; n (%)[b]	
Non-Hispanic White	11(37.9)
African American	1(3.4)
Hispanic/Latinx/Spanish Origin	15(51.7)
Asian	1(3.4)
Middle Eastern	1(3.4)
First-Generation College Student; n (%)[a]	
Yes	23(79.3)
No	6(20.7)
English as First Language; n (%)	
Yes	18(62.1)
No	11(37.9)
Prior Online Course Experience; n (%)	
Yes	27(93.1)
No	2(6.9)

Note. [a] Neither parents or guardians possess a college degree; [b] Percentage does not equal 100% due to rounding

Effectiveness

Effectiveness results for the TAM scales of the sample are presented in Table 3. Results indicated that EpiBot was perceived to benefit course grades for most students (65.5%). Eight participants reported perceived improvements of less than one letter grade (27.6%), eight participants reported perceived improvements of one full letter grade (27.6%), three participants reported improvement by two or more letter grades (10.3%), and 10 students reported their grade would have been the same without using EpiBot (34.5%). Over the course of the semester, a total of 1607 requests (questions) were recorded in EpiBot. EpiBot was unable to find an answer for 46 requests and incorrectly answered 56 requests. Overall, EpiBot correctly answered 93.7% of user questions.

Adoption

The sample (n = 29) submitted a total 1607 requests for an average of 55.4 requests per student. Exact frequency of use per student could not be identified due to the anonymous nature of user requests. As a proxy, students were asked to indicate how often they used EpiBot during the semester. Students reported using EpiBot 6-10 times (41.4%), 1-5 times (20.7%), 11-15 times (17.2%), more than 20 times (13.8%), and 16-20 times (6.9%). Of the 1607 total requests, 1439 (89.5%) were related to epidemiology content corresponding to the 10 course modules and 168 (10.5%) requests were related to course administration (syllabus, due dates, course policies, etc). When asked which platform was used to access EpiBot, participants reported primarily using a computer and mobile device equally (37.8%), primarily a computer (31.1%), and primarily a mobile device (31.1%). When asked which method was used first when having a course- or content-related question, 16 (55.2%) indicated asking EpiBot, 10 (34.5%) indicated reading the syllabus or textbook, 2 (6.9%) indicated using a search engine (e.g. Google), and 1 (3.4%) indicated other with a response of using EpiBot in addition to emailing the instructor. None of the participants indicated they exclusively contacted the instructor or asked a classmate/peer first. Scale scores for TAM-PV, TAM-PaEx, and TAM-PC are presented in Table 3. Three participants reported experiencing barriers/challenges when using EpiBot, which included difficulties accessing the chatbot platform updates, the chatbot's inability to answer some questions, and the need for a stable internet connection to access the platform.

Table 3

Effectiveness and adoption results from Modified Technology Acceptance Model (TAM) scales (n = 29)

Variables	*Scale Range*	*Mean(SD)*
Effectiveness		
TAM-PE	7-35	32.2(3.5)
TAM-PU	7-35	32.3(4.1)
TAM-AT	4-20	14.4(2.8)
TAM-BI	4-20	17.3(3.1)
TAM-SE	2-10	9.1(1.2)
Adoption (individual)		
TAM-PV	3-15	13.2(2.1)
TAM-PaEp	4-20	16.2(2.4)
TAM-PC	3-15	6.5(2.9)

Notes. TAM-AT = Attitudes Toward Chatbot Scale; TAM-BI = Behavioral Intention Scale; TAM-PaEp = Past Experience Scale; TAM-PC = Perceived Complexity Scale; TAM-PE = Perceived Ease of Use; TAM-PU = Perceived Usefulness Scale; TAM-PV = Perceived Voluntariness Scale; TAM-SE = Self-Efficacy Scale

Implementation

The initial iteration of EpiBot was programmed with 141 responses (answers) and 583 intents. To address missing or incorrect responses encountered during the semester, 10 new responses were added as well as 45 intents for a total of 141 answers and 583 intents in the final iteration of EpiBot. Developers identified inconsistent functionality of the voice input feature during the semester. Students were notified and encouraged to use text if the voice feature did not work properly. An initial barrier was a 30-day limit to data storage which required developers to download usage data frequently for record keeping. This limitation was associated with the technology of the platform at the time of EpiBot's development, and data storage limitation were extended during a later AWS platform update.

Time and cost were also considered for EpiBot implementation. EpiBot development and iteration was a time intensive process with an estimate of 50-70 hours required to program and test the initial iteration of the chatbot. EpiBot modification and adaptation was less time intensive due to the interactive dashboard that allows for expedient editing. During the pilot, the instructor reported spending less than an hour per week on modifications and estimated an approximate time requirement of one hour to update course information (due dates, policies, etc) between semesters. The pricing used for the AWS Services integrated into EpiBot are structured as a pay-as-you-go model. Developers pay for the amount and duration of services utilized, with no additional costs of termination fees. Additionally, AWS Free Tier allows developers to use the services for free (up to specified limits), for one year from the data of account creation. Based on AWS Services and data utilized during the pilot, the estimated monthly cost following the year free tier period was $38 a month.

Maintenance

Maintenance at the program, or course-level, was defined as the potential for continued implementation (sustainability) of EpiBot within the epidemiology course. The PSAT domains were used as a framework to structure faculty/developer feedback and facilitate discussion. Strengths that supported maintenance included capacity for adaptation (monthly data review, expedient modification, and scalability); a supportive internal and environmental climate facilitated through partnerships between the course instructor, Computer Science Club, Office of Information Technology, and Office of Academic Technology. Organizational capacity was evidenced through access to EpiBot and related resources through a dedicated website as well as in the student LMS, Blackboard and by continued implementation of EpiBot 24 hours a day, seven days a week. Potential threats to sustaining EpiBot include the need for continued funding in order to support the operating costs of AWS Services and potential organizational capacity issues needed to train new implementers.

Discussion

The purpose of this study was to evaluate the acceptability, feasibility, and effectiveness of a pilot implementation of an epidemiology virtual assistant chatbot (EpiBot) in an online epidemiology course. Using the RE-AIM framework, the findings from this pilot indicated the implementation of EpiBot facilitated student learning and was frequently utilized by participating students throughout the duration of the course. EpiBot was consistently utilized and perceived positively by participants, including from groups that may traditionally achieve lower outcomes and experience more barriers such as first-generation college students, minorities, and those who have low levels of digital literacy (Cataldi et al., 2018; Kaupp, 2012; Tang & Chaw, 2016).

This study offers a novel examination of domain-specific chatbot integration within higher education, focusing on a sample composed predominantly of Hispanic and first-generation college students—a unique demographic that is underrepresented in existing literature. Consistent with recent reviews that report positive findings on chatbot implementation (Kuhail et al., 2023; Nee et al., 2023; Okonkwo & Ade-Ibijola, 2021; Wollny et al., 2021), this study corroborates previously reported beneficial impacts on student learning and satisfaction, despite notable differences in population, domain, and other chatbot characteristics. Concentrating on an online public health epidemiology course, this study mirrors the predominant implementation modality, yet broadens the discussion to an underexplored subject domain. Additionally, EpiBot's focus as a teaching aid and utilization of a predetermined conversational path framework, along with its emphasis on mobile use, underscores the study’s findings on the feasibility and effectiveness of chatbots in specific contexts. This contrasts with chatbots that utilize free or adaptive conversation paths, focus on web-only integration, or purposes aligned to administration, advising, or research.

The majority of user intents corresponded to epidemiologic content, with a smaller proportion addressing administrative course-related information. With a total of 1607 requests from the sample over the duration of the semester, this pilot provides further support that the use of a virtual assistant in an online course may reduce the instructor burden of answering frequently asked questions, such as course policies, and allow for more time to be devoted to more complex student questions (Goel & Polepeddi, 2018). While it is unlikely that all requests would have been asked of the instructor in the absence of the chatbot, it is possible and likely that access to the chatbot may have reduced anxiety and apprehension among students who may be hesitant to ask course-related and content questions (Ayedoun et al., 2015). It is also possible that the novelty of

EpiBot resulted in student curiosity, the development of intrinsic motivation, and a sense of control in the course which led to continued chatbot use (Oudeyer et al., 2016). Consistent with previous chatbot literature, students and the course faculty indicated that chatbot did not replace the role of the instructor but provided an alternative source of information that facilitated learning (Winkler & Söllner, 2018).

Examination of EpiBot use indicated consistent trends of use during the semester, with peak spikes during the midpoint and end of the semester which corresponded to the midterm and final examinations. Use of EpiBot was also facilitated by 24/7 access to the tool and data evidenced a large proportion of requests were made when instructors may not be readily available, such as evenings and weekends. Responses from the pilot indicated participants were comfortable accessing EpiBot on mobile devices which is consistent with previous literature supporting positive perceptions and use of mobile learning in higher education (Crompton & Burke, 2018). Providing students access to EpiBot on mobile devices may have increased use due to increased flexibility and fewer temporal or spatial limitations (Joo et al., 2016).

Overall, EpiBot accurately answered 93.7% of user requests correctly and only three participants indicated incorrect responses as a barrier of using EpiBot during the semester. It should also be noted that a significant number of incorrect responses provided to users were due to students attempting to copy and paste quiz or exam questions directly into EpiBot. There was an intentional choice when designing the course and virtual assistant to require students to demonstrate application, synthesis, and evaluation of concepts which has been demonstrated to increase student learning outcomes (Adams, 2015). The overall accuracy of EpiBot when responding to user inputs can be attributed to the relative extensiveness of the intents and responses related to course content as well as the functionality of the natural language processing of the AWS platform. Previous literature has also indicated the importance of a comprehensive response system in order to reduce barriers to use and decrease frustration that can be caused by continual incorrect responses (Allison, 2012).

Participants in the course perceived EpiBot to be useful tools that made it easier to study course content, accomplish course tasks more quickly, increase productivity, and enhance course effectiveness. Additionally, participants indicated that the use of EpiBot facilitated learning and resulted in improved course performance, with over 72% reporting a positive effect on course grades as a result of the chatbot while no student reported an adverse effect on course grades. These findings are consistent with previous research indicating beneficial student outcomes as a result of chatbot use in education (Abbasi & Kazi, 2014; Goel & Polepeddi, 2018; Kuhail et al., 2023; Nee et al., 2023). Unlike pedagogical agents which interface between the leaner and content or intelligent tutoring systems that provide immediate instruction or feedback to learners (Ayedoun et al., 2015; Baker, 2016; Govindasamy, 2014), EpiBot has the potential to interact with students synchronously in order to respond to individual learner intents. This autonomy can improve outcomes by allowing students to exhibit active control over their learning which aligned with the principles of self-determination and constructivist learning theories (Glasgow et al., 1999; Ryan & Deci, 2000).

Participants also perceived EpiBot as easy to use and indicated positive attitudes towards its use. This is critical as virtual assistant chatbots that do not provide access and are not user friendly will result in decreased value and use from participants (Gupta & Bostrom, 2013). While participants in the pilot indicated having prior experience with using the internet and search engines, the majority did not have prior experience using chatbots but still reported high self-efficacy to use EpiBot. These findings are important to note as it has been demonstrated that participants who exhibit a positive attitude towards chatbots and possess high self-efficacy observe positive effects on chatbot-mediated learning process quality and outcomes (Söllner et al., 2018). Acceptability was also demonstrated through the use of EpiBot as the preferential method of finding information and answers to questions as opposed to other available search engines such as Google.

Finally, program evaluation using the RE-AIM and PSAT frameworks by program stakeholders identified both strengths and challenges that can be used to improve future implementation and adoption. Importantly, this pilot provided evidence that EpiBot, consistent with other chatbots in education, can facilitate student learning through support that requires limited financial and organizational resources once the bot and data have been developed (Winkler & Söllner, 2018). Following development, EpiBot represents a cost-effective and scalable approach to improving epidemiology content knowledge and skills and can extend beyond the university course. However, it is necessary to consider the challenge and time necessary to develop an extensive range of inputs and intents that can distinguish user requests, provide accuracy of response, and keep student frustration low. Chatbot programming can be intensive, requiring interdisciplinary collaboration for required technical expertise in coding and programming and epidemiology expertise to build the content within the chatbot. Expanding collaboration and contributions from faculty, students, and external experts could assist in developing a more expansive and comprehensive knowledge base. An additional strength of EpiBot that was identified in the evaluation was the 24/7 accessibility via a computer or mobile device and the included tutorial video that was provided to participants in the course. It has been previously reported that a lack of direct accessibility and instruction on using the chatbot could result in unsuccessful implementation (Söllner et al., 2018).

Study Limitations

The results of this study should be understood within the context of the associated limitations. This was a pilot study without participant randomization to a control condition which limits the assertion that results were attributed to the chatbot alone. Results may also not be generalizable to other settings due to sample characteristics that were predominantly female, Hispanic/Latinx, first-generation students, with prior online course experience. While students reported EpiBot as a useful tool to increase academic performance in the course, there were self-reported perceptions that did not assess objective measures of temporal trends on knowledge and skill retention. Future research should aim to elucidate effects on objective learning outcomes.

Conclusion

In conclusion, this study's results indicate that implementing a virtual assistant chatbot, EpiBot, in in an online epidemiology course was well received by participants and was perceived as beneficial in increasing course performance. Stakeholder review also identified several advantages associated with EpiBot implementation including cost-effectiveness, flexibility, and scalability. However, further investigation is necessary to measure objective and sustained learning outcomes among participants and to definitively attribute the intervention benefits of EpiBot through the addition of a control group. Overall, this pilot study contributed valuable insights into the potential of virtual assistant chatbots to enhance the online learning experience.

References

Abbasi, S., & Kazi, H. (2014). Measuring effectiveness of learning chatbot systems on student’s learning outcome and memory retention. *Asian Journal of Applied Science and Engineering*, *3*(2), 251-260.

Adams, N. E. (2015). Bloom’s taxonomy of cognitive learning objectives. *Journal of the Medical Library Association: JMLA*, *103*(3), 152.

Allison, D. (2012). Chatbots in the library: is it time? *Library Hi Tech*, *30*(1), 95-107.

Alqurashi, E. (2019). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education*, *40*(1), 133-148.

Anderson, T. (2008). *The theory and practice of online learning*. Athabasca University Press.

Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science*, *60*, 1433-1442.

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, *26*, 600-614.

Bawa, P. (2016). Retention in online courses: Exploring issues and solutions—A literature review. *Sage Open*, *6*(1).

Bell, B. S., & Federman, J. E. (2013). E-learning in postsecondary education. *The Future of Children*, 165-185.

Berzosa, A., Bernaldo, M., & Fernández-Sanchez, G. (2017). Sustainability assessment tools for higher education: An empirical comparative analysis. *Journal of Cleaner Production*, *161*, 812-820.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1-13.

Calhoun, A., Mainor, A., Moreland-Russell, S., Maier, R. C., Brossart, L., & Luke, D. A. (2014). Peer reviewed: using the program sustainability assessment tool to assess and plan for sustainability. *Preventing Chronic Disease*, *11*.

Cameron, M., Lacy, T. A., Siegel, P., Wu, J., Wilson, A., Johnson, R., Burns, R., & Wine, J. (2021). *2019-20 National Postsecondary Student Aid Study (NPSAS: 20): First Look at the Impact of the Coronavirus (COVID-19) Pandemic on Undergraduate Student Enrollment, Housing, and Finances (Preliminary Data).* NCES 2021-456. National Center for Education Statistics.

Cataldi, E. M., Bennet C. T., & Chen, X. (2018). First-generation students college access, persistence, and postbachelor's outcomes (NCES 2018-421). U.S. Department of Education, National Center for Education Statistics.

Cavanaugh, J., & Jacquemin, S. J. (2015). A large sample comparison of grade based student learning outcomes in online vs. face-to-face courses. *Online Learning*, *19*(2).

Centers for Disease Control and Prevention. (2011). *Principles of epidemiology in public health practice: An introduction to applied epidemiology and biostatistics* (3rd ed.). Atlanta, GA.

Crompton, H., & Burke, D. (2018). The use of mobile learning in higher education: A systematic review. *Computers & Education*, *123*, 53-64.

Crutzen, R., Peters, G.-J. Y., Portugal, S. D., Fisser, E. M., & Grolleman, J. J. (2011). An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study. *Journal of Adolescent Health*, *48*(5), 514-519.

Cunningham-Nelson, S., Boles, W., Trouton, L., & Margerison, E. (2019). *A review of chatbots in education: practical steps forward. 30th annual conference for the australasian association for engineering education (AAEE 2019): educators becoming agents of change: innovate, integrate, motivate* (pp. 299-306). Engineers Australia.

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, *24*(4), 38-42.

Gardner, C., & Amoroso, D. L. (2004). *Development of an instrument to measure the acceptance of internet technology by consumers*. In 37th Annual Hawaii International Conference on System Sciences, *2004.* IEEE,

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *American Journal of Public Health*, *89*(9), 1322-1327.

Goel, A. K., & Polepeddi, L. (2018). Jill Watson. *Learning engineering for online education: Theoretical contexts and design-based examples*. Routledge.

Govindasamy, M. K. (2014). Animated Pedagogical Agent: A Review of Agent Technology Software in Electronic Learning Environment. *Journal of Educational Multimedia and Hypermedia*, *23*(2), 163-188.

Gupta, S., & Bostrom, R. (2013). Research note—An investigation of the appropriation of technology-mediated training methods incorporating enactive and collaborative learning. *Information Systems Research*, *24*(2), 454-469.

Hiremath, G., Hajare, A., Bhosale, P., Nanaware, R., & Wagh, K. (2018). Chatbot for education system. *International Journal of Advance Research, Ideas and Innovations in Technology*, *4*(3), 37-43.

Hrastinski, S. (2008). Asynchronous and synchronous e-learning. *Educause Quarterly*, *31*(4), 51-55.

Hung, M.-L., Chou, C., Chen, C.-H., & Own, Z.-Y. (2010). Learner readiness for online learning: Scale development and student perceptions. *Computers & Education*, *55*(3), 1080-1090.

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers & Education: Artificial Intelligence, 1*, 100001.

Joo, Y. J., Kim, N., & Kim, N. H. (2016). Factors predicting online university students' use of a mobile learning management system (m-LMS). *Educational Technology Research and Development*, *64*, 611-630.

Kaupp, R. (2012). Online penalty: The impact of online instruction on the Latino-White achievement gap. *Journal of Applied Research in the Community College*, *19*(2), 3-11.

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, *28*(1), 973-1018.

Luke, D. A., Calhoun, A., Robichaux, C. B., Elliott, M. B., & Moreland-Russell, S. (2014). Peer reviewed: the program sustainability assessment tool: a new instrument for public health programs. *Preventing Chronic Disease*, *11*.

Ma, P. W., & Shea, M. (2021). First-generation college students' perceived barriers and career outcome expectations: Exploring contextual and cognitive factors. *Journal of Career Development*, *48*(2), 91-104.

Martin, F., Budhrani, K., & Wang, C. (2019). Examining faculty perception of their readiness to teach online. *Online Learning*, *23*(3), 97-119.

Martin, F., Wang, C., & Sadaf, A. (2018). Student perception of helpfulness of facilitation strategies that enhance instructor presence, connectedness, engagement and learning in online courses. *The Internet and Higher Education*, *37*, 52-65.

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, *115*(3), 1-47.

Mikic-Fonte, F. A., Llamas-Nistal, M., & Caeiro-Rodríguez, M. (2018). *Using a chatterbot as a FAQ assistant in a course about computers architecture*. 2018 IEEE Frontiers in Education Conference (FIE),

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education*, *14*(2), 129-135.

National Center for Education Statistics. (2023). *Undergraduate Enrollment. Condition of Education.* U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/cha.

Nee, C. K., Rahman, M. H. A., Yahaya, N., Ibrahim, N. H., Razak, R. A., & Sugino, C. (2023). Exploring the trend and potential distribution of chatbot in education: A systematic review. *International Journal of Information and Education Technology*, *13*(3), 516-525

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, *2*, 100033.

Oudeyer, P.-Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Progress in Brain Research*, *229*, 257-284.

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Journal of Educational Technology & Society*, *12*(3), 150-162.

Ritzhaupt, A., Poling, N., Frey, C., & Johnson, M. (2014). A synthesis on digital games in education: What the research literature says from 2000 to 2010. *Journal of Interactive Learning Research*, *25*(2), 261-280.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Sinha, S., Basak, S., Dey, Y., & Mondal, A. (2020). *An educational chatbot for answering queries. Emerging Technology in Modelling and Graphics*: Proceedings of IEM Graph 2018,

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, *151*, 103862.

Söllner, M., Bitzer, P., Janson, A., & Leimeister, J. M. (2018). Process is king: Evaluating the performance of technology-mediated learning in vocational software training. *Journal of Information Technology*, *33*, 233-253.

Song, D., Oh, E. Y., & Rice, M. (2017). *Interacting with a conversational agent system for educational purposes in online courses.* In 2017 10th International Conference on Human System Interactions (HSI).

Stoll, S., Janevic, M., Lara, M., Ramos-Valencia, G., Stephens, T. B., Persky, V., Uyeda, K., Ohadike, Y., & Malveaux, F. (2015). Peer reviewed: A mixed-method application of the program sustainability assessment tool to evaluate the sustainability of 4 pediatric asthma care coordination programs. *Preventing Chronic Disease*, *12*.

Strahan, B., Calhoun, J., & Potterveld, B. (2017). *Create a question and answer bot with Amazon Lex and Amazon Alexa*. https://aws.amazon.com/blogs/machine-learning/creating-a-question-and-answer-bot-with-amazon-lex-and-amazon-alexa/

Sun, A., & Chen, X. (2016). Online education and its effective practice: A research review. *Journal of Information Technology Education*, *15*.

Tang, C. M., & Chaw, L. Y. (2016). Digital Literacy: A Prerequisite for Effective Learning in a Blended Learning Environment? *Electronic Journal of E-Learning*, *14*(1), 54-65.

Thomas, H. (2020). Critical literature review on chatbots in education. *IJTSRD*, *4*(6), 786-788.

Verleger, M., & Pembridge, J. (2018). *A pilot study integrating an AI-driven chatbot in an introductory programming course*. 2018 IEEE Frontiers in Education Conference (FIE),

Winkler, R., & Söllner, M. (2018). *Unleashing the potential of chatbots in education: A state-of-the-art analysis.* Academy of Management Proceedings

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet?-a systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, *4*, 654924.

Zhang, W., Wang, Y., Yang, L., & Wang, C. (2020). *Suspending classes without stopping learning: China's education emergency management policy in the COVID-19 outbreak*. In (Vol. 13, pp. 55): MDPI.

Authors Bios

Eric J. Conrad is an Assistant Professor in the Department of Kinesiology and Public Health Promotion at California State University, Stanislaus.
Email: econrad@csustan.edu

Austin White is an undergraduate student in the Department of Computer Science at California State University, Stanislaus

Job M. Ramirez is a Master of Public Health Candidate at the University of Alabama at Birmingham.

Alyssa Alvara is an undergraduate student in the Department of Kinesiology and Public Health Promotion

Laura K. Rees is an MA student in Educational Studies at UMASS Global