Teaching communication skills with technology: Creating a virtual patient for medical students

Baris Sezer
Hacettepe University, Turkey

Tufan Asli Sezer
Ankara University, Turkey

Virtual patient (VP) is a concept used in the teaching of communication skills, and like physical examinations and other professional skills, must be taught with utmost care. In Turkey, as VPs are yet to be used in medical training, the usual practice when teaching such skills is to use standardised patients (individuals pretending to be patients). The main purpose of this study was to design, develop and evaluate a 3D VP application that can move, has speech-over lip sync, allows written communication and is supported by a strong scenario to improve the communication skills of students. The study was designed and carried out using developmental research methods. The implementation phase involved a pretest posttest quasi-experimental design. The participants in the study consisted of academics specialising in medicine, software experts, an education technology expert, an assessment and evaluation expert, and medical students. The study found that VP applications were accepted by students and were as effective as standardised patients for the teaching of communication skills. The students reported that the VP application developed was very successful in terms of visual and behavioural reality.

Introduction

Health care professionals need to possess certain skills (including cognitive, affective, and psychomotor) if they are to ensure patient safety, prevent medical errors, and resolve problems in the most efficient and appropriate manner (Lampotang et al., 2013; MacLean, Kelly, Geddes, & Della, 2017; Oh, Jeon, & Koh, 2015). Psychomotor skills are based on such as hand-eye coordination, cognitive skills involve thinking and involve decision making, and affective skills are the interacting skills, and include interpersonal skills (Lampotang et al., 2013). Knowledge, problem solving, communication skills, and physical examination are the four main components of a good clinical practice (Silverman, Kurtz, & Draper, 2013). Of these, communication skills form the basis of clinical practice, as if this skill is not properly acquired, the other components may become meaningless, because effective physician-patient communication is important in the delivery of high-quality health care (Ha & Longnecker, 2010).

Communication skills are not a personality trait, but rather a series of skills that can be learned. A previous study found that a physician meets with some 200,000 patients throughout their professional career and that serious problems may be encountered in physician-patient interactions, but that improvements in communication skills can make a significant difference in the application of health services (Silverman et al., 2013). Establishing an effective line of communication with patients, patients’ families, and other health care professionals is a prerequisite for the development of a good relationship, and so the teaching of communication skills in medical education cannot be seen as an optional activity (Burke, LeBlanc, & Henneman, 2016). Such skills must be taught with utmost care, and with similar concerns as those required for physical examinations and other professional skills (Buresh & Gordon, 2013; MacLean et al., 2017).

There are a limited number of programs offered by faculties of health sciences in Turkey providing communication skills training in a structured and applied manner. Communication skills training must be applied, as it is necessary to teach attitude as well as skills (Banerjee et al., 2017; Naughton et al., 2018). Conventional master-apprentice style training or didactical teaching methods are ineffective methods for the teaching of communication skills (Silverman et al., 2013). In this context, when planning a communication skills training program, different approaches, such as case-based methods, role-playing or demonstrations can be used to introduce, explain or set out a problem (MacLean, et al., 2017; Oh et al., 2015; Yuan, Williams, Fang, & Ye, 2012), and one of the most important approaches at this stage is simulation (Cook & Triola, 2009; Nehring & Lashley, 2009; Norman, 2012).
Simulation

Simulation was defined by Gaba (2004) as “an educational strategy that replaces or amplifies experiences that replicate aspects of the real world in an interactive fashion” (p. 2). As Figure 1 shows in the Simulation Triangle, there are three main types of simulation (physical, virtual and biologic) used in health care education. The choice is dependent on the skill to be taught (Lampotang et al., 2013).

![Skills Triangle](image1.png)

**Figure 1.** The skills triangle and the simulation triangle in healthcare (Lampotang et al., 2013)

**Physical simulators** used for the teaching of psychomotor skills have no virtual components (Aggarwal et al., 2007). Task trainers used in the teaching of intramuscular injections are examples of this type of simulation. **Virtual simulators** used in the teaching cognitive skills are usually web/screen based, and the Virtual Anesthesia Machine (VAM) developed by Fischler, Kaschub, Lizdas, and Lampotang (2008) and Computer-based Trainer (CBT) developed by Lampotang et al. (2013) stand as good examples of this type of simulation. Finally, **biological simulators** used in the teaching of affective skills (communication skills, collaboration, critical reasoning, self-evaluation, empathy, etc.), provide a highly valid experience to learners, and are usually based on interpersonal interactions (MacLean et al., 2017). Training using standardised patients (SP), that is to say, individuals pretending to be patients, would be an example of this category. SPs are “real patients or healthy individuals trained to simulate the full set of symptoms associated with a condition in a consistent manner, and able to evaluate student performance” (Sarmasoglu, Dinc, & Elcin, 2016, p. 3). SPs are frequently used in faculties of health sciences for communication skills training, and in recent years, virtual patient (VPs) have also started to be used in the teaching of these skills (Cordar, Wendling, White, Lampotang, & Lok, 2017; Oh et al., 2015).

**Standardised and virtual patients**

**Mixed reality humans** are a combination of the biological and virtual simulations noted on skills and simulation triangles (Figure 1), and are based on interactions between a user and an avatar acting like a human (Cordar et al., 2017; Kotranza, Lok, Deladisma, Pugh, & Lind, 2009). This type of simulation is usually referred to as a VP in the literature. (Cook, Erwin & Triola, 2010; LeFlore et al., 2012). A VP can be defined as “a computer program that simulates real-life clinical scenarios in which the learner acts as a health care provider obtaining a history and physical examination, and making diagnostic and therapeutic decisions” (Nestel & Bearman, 2014, p. 9). VPs can be considered a subset of virtual humans just as a patient is a subset of humans (Cordar et al., 2015; Robb, Kleinsmith, Cordar, White, Lampotang et al., 2016). They are typically composed of three main components: inputs (mouse, keyboard, etc.), storage (database, scenario, simulation engine, etc.), and outputs (visual, auditory, etc.) (Cendan & Lok, 2012).
A review of the literature shows that SPs are routinely used in the category of biological simulation for the teaching of affective skills (Arnold et al., 2015; Clay et al., 2000; Lorin, Rho, Wisnivesky, & Nierman, 2006; Powell & Turnbull, 2016), although it is argued that both SPs and VPs are very effective in the teaching of these skills (Lin, Travlos, Wadelin, & Vlasses, 2011; Oh et al., 2015; Stevens et al., 2006). Studies conducted to date report that students see both as useful in this process. In other words, they accept them both as effective learning tools (Botezatu, Hult, Tessma, & Fors, 2010; Cant & Cooper, 2010; Deladisma et al., 2007; Lin et al., 2011; Oh et al., 2015; Stroup, 2014).

VPs have been used for conducting medical interviews (Rizzo et al., 2014; Skarbez, Welch, Brooks, & Whitten, 2017; Stevens et al., 2006), empathetic communication skills (Johnsen, 2008), speaking up (Robb, Cordar et al., 2015; Robb et al., 2014), teaming up (Robb, Kleinsmith, Cordar, White, Lampotang et al., 2016; Robb, White et al., 2015), how medical professionals gaze at VHs (Robb, Kleinsmith, Cordar, White, Wendling et al., 2016), team communication skills (Cordar et al., 2015; Cordar et al., 2017), team training-handoffs (White et al., 2015), diagnostic reasoning skills (Kotranza et al., 2009; Sia, Halan & Lok, 2016), empathy (Kleinsmith, Rivera-Gutierrez, Finney, Cenden, & Lok, 2015), negotiation (Gratch, DeVault, Lucas, & Marsella, 2015), and leadership skills (Campbell et al., 2011). Many studies noted that learners behave in a realistic manner when interacting with these applications, and are highly satisfied with them (Chuah, 2013; Cordar et al., 2015; Johnsen, 2008; Kotranza et al., 2009; 2016; Rizzo et al., 2014; Robb, Kleinsmith, Cordar, White, Wendling et al., 2016; Robb et al., 2014; Sia et al., White et al., 2015). In a study conducted by Johnsen (2008), experts observed and evaluated medical and pharmacy students as they communicated with screen-based VPs, and noted that during the interviews, the students empathised, interacted, acted natural, and were as satisfied with the VPs as they were with SPs. Similar findings have been reported in other studies (Rizzo et al., 2014; Robb et al., 2014; Robb, Cordar et al., 2015).

**Objective of the study**

A review of the international literature shows that VPs are widely used in the field of health sciences, although the concept is relatively new in Turkey. The VP concept remains theoretical in Turkey and is yet to be used in training, although recommendations have been made for its development and use (Mıdık & Kartal, 2015). The authors of this study spent 4 months in the University of Florida, College of Medicine, where virtual training applications are in common and effective use in teaching, and had the opportunity to experience different virtual human applications. Based on their experience, the authors aimed to demonstrate the feasibility of VP teaching methods in Turkey for the teaching of communication skills to medical students. In this regard, the main purpose of the study was to design, develop, and evaluate a VP application supported by a strong scenario to be used by medical students who, for many years, have been taught communication skills using SPs. Every stage in the study is explained in detail in the hope that it can support future researchers in their efforts to develop similar applications. The strengths and limitations of VPs when compared to SPs are also discussed.

**Methods**

This study was designed and carried out using developmental research methods, involving the systematic study of the design, development, and evaluation of instructional programs, processes, and products. These studies usually make use of a generic development cycle, and the application in the present study was developed within the framework of such a basic development cycle (analyses, design, development, implementation, evaluation), with the findings presented accordingly (Kuzu, Çankaya & Mısırlı, 2011; Richey & Klein, 2005). To demonstrate the effectiveness of the developed VP application, a comparison was made with the SP interview through a quasi-experiment involving a pre-test post-test control group design in the implementation phase (Figure 2).
Figure 2. Experimental design of the study

Participants

The study was conducted with 92 participants (their roles will be clarified in data collection section of the paper), with an experiment group that comprised academics specialising in medicine, software experts, an education technology expert, an assessment and evaluation expert, and medical students. The students who made up the experimental and control groups in the implementation phase of the study had no prior experience with VP applications, but all had experience with SP interviews over the past 2 years. An academic achievement test was administered to the students in both groups (experimental group: 44 students; control group: 44 students) as pre-test. An independent samples \( t \)-test (Table 1) was conducted to analyse their pre-test scores, showed there were no significant differences between the two groups (\( t(42) = .079, p > 0.05 \)).

Table 1

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>X</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>44</td>
<td>57.09</td>
<td>11.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>44</td>
<td>56.81</td>
<td>11.33</td>
<td>.079</td>
<td>42</td>
<td>.826</td>
</tr>
</tbody>
</table>

At the beginning of the semester the students were randomly assigned to small group sessions for the development of communication skills. Equal numbers of students assigned to each group based on gender, age, whether their program offers instruction in Turkish or in English, and whether they were repeat students or not. Because there were no significant differences between their academic achievement pre-test scores, the two groups were assumed to be comparable. Informed consent was obtained from each research participants. The students were informed that the test would in no way affect their grades for the semester. Table 2 presents demographic information about the students.
Table 2
Demographics of participants

<table>
<thead>
<tr>
<th></th>
<th>Experimental</th>
<th></th>
<th>Control</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Gender:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24</td>
<td>54</td>
<td>22</td>
<td>50</td>
<td>44</td>
<td>52</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td>46</td>
<td>22</td>
<td>50</td>
<td>42</td>
<td>48</td>
</tr>
<tr>
<td>Age:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 - 21</td>
<td>36</td>
<td>81</td>
<td>38</td>
<td>86</td>
<td>74</td>
<td>84</td>
</tr>
<tr>
<td>22 - 25</td>
<td>8</td>
<td>19</td>
<td>6</td>
<td>14</td>
<td>14</td>
<td>16</td>
</tr>
</tbody>
</table>

Data collection and analysis

At every stage of the study data was collected for different purposes. The process of data collection was as follows:

Phase 1 - Analysis: In this stage, the expert opinion and students’ views were collected using interview forms in order to identify the requirements for the study.

Phase 2 - Design: In this stage, VP applications developed by the Virtual Experiences Research Group (VERG) of Florida University were examined, and recommendations were collected regarding which of them could be of potential use in the present study.

Phase 3 - Development: In this stage, evaluation forms were used to obtain feedback from the students and experts on the first version of the VP application, and the system logs were evaluated.

Phases 4 and 5 - Evaluation and Implementation: In this final phase, the usage logs generated by the system following the use of the VP application by students of Hacettepe University’s Faculty of Medicine were evaluated, their cognitive gains were measured using the achievement test presented below, and opinions about the application were garnered using the interview form.

Achievement test
The achievement test was developed by the researchers and consisted of 10 items, containing questions of many different types (multiple choice, fill-in-the-blank, and open-ended questions). A table of specifications was prepared and used to ensure content validity. Care was taken to make sure that each question was equally valid for the scenarios used with students who communicated with the VP and students who communicated with the SP. In addition, prior to administering the test, evaluation criteria were developed for the achievement test, and field specialists and students of the Faculty of Medicine were consulted to establish validity and reliability of the test. The achievement test was administered to both groups as a pre-test one week prior to the physician-patient interviews, and as a post-test immediately after interviews with the VP or SP. Each item on the test was scored out of 10 points, with the maximum total score on the test being 100 points. The test was scored by two medical instructors and one measurement and evaluation specialist. The analytic rubric was used to score the students’ pretest and posttest.

Semi-structured interview form
The interview form, consisting of three items, was developed by the researchers, and was used to gain a deeper insight into data garnered via the VP application. The interviews conducted with 41 students on the on the day after they had used the VP application were intended to gain negative and positive feedback on the application and areas of possible improvement, as well as whether the respondents would be open to using the VP application in the future for communication skills training.

Data analysis
The data collected in the study was analysed using qualitative and quantitative methods. A descriptive analysis was used to analyse qualitative data, while descriptive statistics were used to analyse the quantitative data collected at the different stages of the study. The statistical analysis of the quantitative data was conducted using the Statistical Package for the Social Sciences (SPSS) 22.0 software package. The collected data was coded qualitatively, and to ensure reliability, the students’ opinions were analysed separately by the two researchers, and the level of agreement was found to be 0.94 based on the following
formula: number of agreements/(total number of agreements + disagreements) x 100 (Miles & Huberman, 1994).

Findings

The findings of the study are presented here in stages corresponding to the steps in a basic development cycle.

Analysis phase

In the analysis phase of the study, literature was reviewed and the findings of various studies on SPs and VPs were examined. As Table 3 shows, there are many advantages to using a VP. The literature review also showed that the students’ acceptance of VPs increased in direct proportion to visual reality (Kotranza et al., 2009; Robb, White et al., 2015) (i.e., animation quality, lip sync, image quality, etc.) and behavioural reality (Robb, Cordar et al., 2015) (i.e., audio quality, eye contact, appropriate behaviour, non-verbal behaviours, etc.).

Table 3
Comparing SPs and VPs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Standardised patient</th>
<th>Virtual patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency</td>
<td>Variable due to unavoidable human factors.</td>
<td>Consistent, always gives the same responses without memory lapses, mistakes, etc.</td>
</tr>
<tr>
<td>Adaptability</td>
<td>Requires practice. Adaptability is possible after some practice. However, some roles are difficult to play.</td>
<td>Can be programmed for all scenarios, given time.</td>
</tr>
<tr>
<td>Manpower</td>
<td>A SP is an individual person.</td>
<td>A single VP can interact with multiple students at the same time.</td>
</tr>
<tr>
<td>Costs</td>
<td>SPs are paid USD20 per hour in other countries. In Turkey, they are paid at least TRY70 (about USD10) for 2 to 3 hours of work.</td>
<td>After the initial cost there are no other costs, or very small costs required for small adjustments.</td>
</tr>
<tr>
<td>Versatility</td>
<td>The ability to play characters of different genders, ages, races, body shapes, voices, accents, etc. is limited.</td>
<td>Can play any character without limitation.</td>
</tr>
<tr>
<td>Logistics</td>
<td>Requires institutions to have separate rooms and large waiting areas so that the SP and the students do not encounter one another prior to their interview. In addition, providing group training at the same time can be very difficult due to such issues as holidays, illness, running late, etc.</td>
<td>Can be used or updated 24/7.</td>
</tr>
</tbody>
</table>

A scenario was created based on the pre-defined steps of medical interview stages in patient-physician interaction, as recommended in Dickerson et al. (2005): greeting, chief complaint, history of present illness, other active problems, past medical history, family history, social history, review of symptoms, physical exam, impression and plan, and closing. This led to the creation of the analysis document. In addition, meetings were held with the Center for Safety, Simulation & Advanced Learning Technologies (CSSALT) and Computer, Information Sciences and Engineering Department teams of the Florida University to learn about their design experience and preliminary recommendations. On this basis, the researchers created a screen-based mock-up that allows interactions with written communications, provides feedback to students during and after practice, allows repeated practice, provides feedback about user behaviours, and is multi-platform compatible (mobile, tablet, personal computer etc.). The production inputs (process chart, metrics,
quality criteria, standards, etc.) were created using this model. Then, an analysis document was created and the design phase was initiated.

**Design phase**

The steps taken in the design phase, taking the analysis document as the basis, were as follows:

- Examine the analysis document with field specialists, exchange opinions about the general framework of the application,
- Examine the virtual human applications used in other countries with experts, identify challenges faced in their use, develop recommendations to overcome these challenges,
- Design sub-components (modules) of the application in detail,
- Detail component features,
- Specify the interactions between components,
- Define the system interface,
- Create the virtual character and environment,
- Create the user scenario,
- Design data collection instruments for the collection of user feedback about the system, and
- Create the design document and visualise the system in the form of screen images.

Based on meetings with experts and the literature review, the researchers decided to: (1) prepare Turkish usage functions, (2) prepare a feedback collection function, (3) ensure compatibility with all platforms (tablet, smart phone etc.), (4) create an environment similar to the SP meeting room in Hacettepe University’s Faculty of Medicine in Turkey, where the application will be used, and (5) select a typical Turkish character from the Virtual People Factory (VPF - http://vpf2.cise.ufl.edu) system developed by the VERG team of Florida University. The virtual environment and the VP character were designed on this basis. We used Adobe Fuse™ for creating characters and Lipsync Pro™ for lip synchronising. The whole 3D scene was created in Unity 3D™. The VP was selected as a 55 year old male character (named Mr. Nihat). Figure 3 illustrates the initial version of the Mr. Nihat character.

![Mr. Nihat Character](image)

**Figure 3.** The real (on the left) and virtual (on the right) environments

Following the creation of the clinical environment and the VP (Mr. Nihat), the researchers decided to create a scenario for “the evaluation of a patient suffering from a headache”. To this end, first, expert opinion was obtained, and a scenario document was created on the basis of the feedback received from students in previous years, and from sample communication skills videos used routinely in class. The scenario was then organised in line with the flow specified in the analysis stage, and the basic elements of a medical meeting were set out in the form of questions and answers. The draft scenario was shared with the experts via e-mail and revised on the basis of their recommendations.

As the final step in this stage, apart from the technical aspects explained above, the “Feedback/Interview Form” to be filled out by the students and the “Academic Achievement Test” were created with the input of experts (explained in the section on data collection instruments). The design document was thus created and the researchers moved on to the development phase.
Development phase

The steps taken in the development phase, taking the design document as the basis, were as follows:

- Code every component (character, environment, scenario) in line with the design, and create a product prototype,
- Conduct unit tests and repeat the development and testing cycle until the desired results are obtained, have experts and students use the application in this process and obtain feedback,
- Check the integrity of the components, and
- Carry out confirmation and validation, and develop the first version of the product.

First, to create the product prototype, the environment and the character created were uploaded to the VP factory. Patient information, goals, scope, and stages were defined on the VP factory. Next, the scenario system prepared in the previous phase was created and added in stages (greeting, history of present illness, medical history, social history, family history, and closing).

As shown in Figure 4, different variations were entered into the system for each question the students could ask. For example, variations of the question about age included “How old are you?”, “May I ask your age?”, “Your age?”, “What is your age?” and “Age?” When a learner interacts with the VP, the learner’s input is matched against all the questions in the corpus, and the closest match is found. The corresponding response for the identified match is displayed as the response from the VP. The important questions on each stage were marked as important discoveries (13 important discoveries in total). This critical discoveries were determined by experts. When the students interact with VP, if there is a critical discovery the system informs the students “You made a critical discovery”. This may be evaluated as formative/summative assessment logs for students. It can also be discussed in debriefing sessions. After this step, the technical works started.

![Image of Virtual People Factory 2.0](image)

**Figure 4.** Entering the scenarios into the system in stages

All the VP’s responses were recorded via a microphone to .wav sound files by the first author. The recorded sound files were integrated/uploaded to the VP factory. In line with the revision requests received from the students, a timer showing the interview duration and a note-taking field for use during the interviews were added to the system. The first version of the VP application, shown in Figure 5, was then created and released for the use of students.
Implementation and evaluation phases

The VP application was first piloted with a sample of 18 students, and areas of potential improvement were identified based on logs and student interviews. These included: (1) enriching the help menu (a web interface was developed for the original application, and a help menu was created in the form of sub-components); (2) including a sample VP-physician interview (a sample VP practice video was added to the web interface); and (3) providing a more practical method for feedback (a feedback form was created using Google Form and integrated into the VP application via the VP factory).

The logs from the pilot study were categorised as failure cases, as shown in Figure 6, and the following actions were taken:

A) Entry does not exist (19% of all queries): The logs showed that the students asked questions that were unanticipated by the researchers. For example, they asked whether the patient had previously undergone an operation. To resolve this problem, new questions and answers were added within the limits of the scenario.

B) Variation of query phrasing (9% of all queries): The students used unanticipated variations of questions, for example asking, “Is everything alright?” instead of “How are you feeling?” To resolve this problem, additional variations to the questions were added in the following phase.

C) Joined questions (7% of all queries): Students asked multiple questions at once. For example, they asked, “Are you on medication, and what do you do when you get a headache?” To resolve this problem, a help menu was added to the system and examples were provided of questions that can be asked and question types that would not be appropriate.

D) Declarative statements (2% of all queries): The students sometimes typed declarative statements, such as, “Hello, Mr. Nihat. I am a third-year student. May I ask what brings you here today?” As it would require artificial intelligence for VPs to make sense of such long sentences, an additional explanation was provided in the help menu about the limitations of the VP character.
Following the pilot and the revisions, the experimental group of 44 students practiced with the VP, and the 44 students in the control group practiced with SPs who are routinely used in training. Following the communication skills training provided by the VP and the SP, the scores received by the students on the achievement test conducted after the intervention (post-test) were compared to identify any differences between the academic achievement levels of the experimental and control group students. The steps followed for this comparison and the null hypothesis used was as follows.

H0: There is no significant difference between the academic achievement scores of the experimental and control groups following the intervention.

An independent samples t-test was conducted to test hypothesis H0, and the results are reported in Table 4.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>X</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>44</td>
<td>80.63</td>
<td>9.88</td>
<td>.956</td>
<td>42</td>
<td>.647</td>
</tr>
<tr>
<td>Control</td>
<td>44</td>
<td>77.50</td>
<td>11.79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Both groups received 2 hours of theoretical communication skills training, after which, the first author made a 15-minute presentation to the experimental group describing the VP and its characteristics and explaining its use. The students were then assigned usernames and passwords, and allowed to access the system at their own convenience. Students accessed the system from different places (dorm, home, faculty) and using different platforms (mobile, personal computer, tablet), and after use, the effectiveness of the script-based approach to medical communication skills training was evaluated, the students’ logs were acquired from the VP factory and a focus group meeting was conducted with these students.

The results from the main application after the revision are presented in Table 5, which shows that the mean academic achievement score of the experimental group was 80.63, whereas the mean academic achievement score of the control group was 77.50. The t-test results reported in the table identify no significant difference between the academic achievement post-test scores of the experimental or control groups (t(42) = .956, p > 0.05), and so hypothesis H0 is accepted. In other words, the academic achievement levels of the experimental group students who received communication skills training using the VP were no different to those of the control group students who received communication skills training with the SP. This finding indicates that the VP, when used as part of a good medical practice program, achieved results that were as positive as those achieved with a SP, even though the students had no prior experience with VPs. The logs received showed that many of the revisions made after the pilot study proved useful, and these results - Entry does not exist (8% of all queries), Variation of query phrasing
(4% of all queries), Joined questions (2% of all queries), Declarative Statements (1% of all queries) - should be considered very encouraging.

Another important finding was that 83% of the 41 experimental group students who participated in the semi-structured face-to-face interview responded positively to the question: “Would you like to use other VPs during your medical education?” Table 5 presents details of the feedback received via the interview form regarding the strengths and weaknesses of the VP and areas of potential improvement, as well as the thematic codes based on these findings. The codes were created through an analysis of the process evaluations made by students at the end of the semester.

Table 5
Thematic codes related to the VP application

<table>
<thead>
<tr>
<th>Weaknesses</th>
<th>n</th>
<th>%</th>
<th>Code</th>
<th>n</th>
<th>%</th>
<th>Code</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty of typing</td>
<td>12</td>
<td>48</td>
<td>Repeated practice</td>
<td>32</td>
<td>40</td>
<td>Speech recognition</td>
<td>15</td>
<td>47</td>
</tr>
<tr>
<td>Delays</td>
<td>8</td>
<td>32</td>
<td>Visual reality</td>
<td>18</td>
<td>22.5</td>
<td>Artificial intelligence</td>
<td>10</td>
<td>31</td>
</tr>
<tr>
<td>Inability to recognise typos</td>
<td>5</td>
<td>20</td>
<td>Strong scenario</td>
<td>16</td>
<td>20</td>
<td>Speed</td>
<td>7</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Instant feedback</td>
<td>14</td>
<td>17.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As Table 5 shows, the students reported that the strengths of the simulation included the opportunity for repeated practice (which is arguably its most important feature), animated visuals, 3D, lip sync, a comprehensive scenario, and instant visual feedback provided by the application to students after every question they ask. On the downside, the students thought that the speech communication could be more effective, and criticised the occasional delays in the connection with the server and the inability of the system to automatically correct typos as weaknesses. The students would prefer a speedy system that utilises artificial intelligence, recognises the questions asked and responds accordingly.

Discussion

In different parts of the world, virtual human applications are seeing increasing use in many service fields (banking, telecommunication, etc.) and formal/informal training scenarios, especially in health sciences. Studies show that VPs, a category of virtual humans that have been in use in the field of medicine for the teaching of affective skills since 2006, are accepted by students and are at least as effective as SPs (Cordar et al., 2015; Cordar et al., 2017; Gratch et al., 2015; Johnsen, 2008; Kleinsmith et al., 2015; Rizzo et al., 2014; Robb, Cordar et al., 2015; Robb, Kleinsmith, Cordar, White, Lampotang et al., 2016; Robb, Kleinsmith, Cordar, White, Wendling et al., 2016; Robb et al., 2014; Robb, White et al., 2015; Skarbez et al., 2017; Stevens et al., 2006). The recommendations of the studies regarding the creation and utilisation of a VP were taken up in the relevant phases of the study, and the following three main design elements were incorporated.

1. Visual quality elements (animation quality, 3D visuals, lip sync, image resolution, creation of a character that fits the culture, good fit between the character and the scenario, a virtual environment that resembles the users’ prior experiences) were paid careful attention in the design and development phase.
2. Behavioural reality (sound quality, intonation adjusting depending on the response, VP making eye contact with the user, an animated character) was achieved through various programs (Adobe Fuse, LipSync, Unity 3D) in the development phase.
3. An effective scenario was prepared (the scenario was based on good practice videos currently used in training, expert opinion was consulted at every stage, a pilot study was conducted and revisions were made accordingly, there was a pre-set flow, and users were given feedback in the form of reminders when they missed some questions).

Students in the experimental group had academic achievement levels that were as good as those of the control group students, even though they had no prior VP experience. This may be attributed to a number of factors, among them, the fact that the VP was a novel technological application (animated, lip synced, 3D) and thus motivating for the students, that instant feedback was provided, reminding the students of the
basic questions in communication, and that ability to arrange patient-physician interviews at the students’ own convenience.

A review of literature in Turkey reveals that VPs are yet to be used in Turkey, and this study and application can be expected to pioneer the adoption of VPs in health care education at a national level. In Turkey, SPs are used for the teaching of affective skills, although the use of SPs has its own limitations, which means that different training alternatives should be considered. As Table 3 shows, there are many advantages to using a VP, and the main finding of the present study is that using a VP is as effective as using a SP in terms of improving academic achievement, and should be evaluated with the above considerations in mind. For example, there were zero or minimal ongoing costs associated with the use of a VP, while the SPs employed in the control group activities had to be compensated. As further benefits of VP training, all students received the same feedback/responses, the students were able to practice at their own convenience, and the labour hours used in the control group, in which seven different SPs were used, were not needed. Besides, the students (the Y generation) prefer interacting with the information and communication technologies during learning process (Helsper & Eynon, 2010). At this point, we may use technology to make effective courses according to our course objectives.

The logs for the male and female students who used the application were examined, and no gender related differences were found, although previous studies in literature have reported contrasting findings in this regard (Marei, Donkers, Al-Eraky, & van Merrienboer, 2017). Furthermore, other studies have reported that the higher the cognitive load of students in these applications, the greater the negative effect on their learning experiences (Choi, Lindquist, & Song, 2014). Similarly, it has been reported that students who encounter complicated problem situations or scenarios without prior briefing face a higher cognitive load, and fail to achieve effective learning (Clark, Kirschner, & Sweller, 2012; Kirschner, Sweller, & Clark, 2006; van Merrienboer & Sweller, 2010). Taking these considerations into account, the same instructor provided theoretical training on communication skills in the same lecture hall to all students, while the VP application only covered the questions they were expected/required to ask via a simple interface and help menu.

The positive results obtained in the study can be attributed to the care taken to ensure visual and behavioural reality, as recommended in studies into the use of virtual humans, receiving expert opinion and the students’ views during the development phase, and the creation of the virtual character, virtual environment, and scenario based on this feedback. To enable written communication between the user and the machine, a special scenario tree was created for the headache scenario using a natural dialogue interface. The scenario tree contained paths extending from the main menu to sub menus, with the required concepts, expressions and stages defined for each path. Concepts are structures that are defined to facilitate the matching of user scripts with appropriate branches on the scenario tree. Words that have the same root and that are differentiated only with inflectional suffixes are associated with the same concept. For example, the Turkish words “yapacağım” (“will do”), “yapmak” (to do”) and “yaptık” (“have done”) have the same root and represent the same meaning for the routing system, and as a result, are associated with the concept “yapmak”. Variations were created for all questions based on this branching structure, on the data from the pilot study and on the expert predictions, and all question variations were integrated into the system. In this way, the questions entered by the students were matched with a high level of accuracy, with the responses kept on the database for the VP to provide in the future.

The application developed for this study was first tested via a pilot study, and revisions to the application were made based on data from this pilot study, making the system stronger. A web page was created based on the feedback from the pilot study containing such informative documents as “What is a VP?” and “How to use the VP”, as well as a sample VP practice video. As a result, the problems with script matching encountered in 37% of the cases in the pilot study were reduced to 15% in the main study. In future studies, researchers would be well advised to pay careful attention to this issue.

Reports were generated for the first use of the application by the students. The students who engaged in repeated practice (14 students had repeat meetings) were 22% more effective in their later interviews than in their first interview. It is reported that the smaller this figure, the higher the level of acceptance and satisfaction among users (Cordar et al., 2017; Johnsen, 2008). In this respect, it is recommended that researchers conduct pilot studies and prepare strong help menus in future studies. The structuring of the scenario based on a specific flow (greeting, chief complaint, history of present illness, other active
problems, past medical history, family history, social history, review of symptoms, physical exam, impression, and plan), and the ability of the system to provide students with instant and visual feedback at every interaction were identified as the strengths of the application.

Conclusion

The medical students who took part in the study reported that the VP with script-based matching performed well. Although they have their own limitations, VPs were found to be at least as effective as SPs based on the advantages they offer, and should therefore be considered for use in affective skills training.

Limitations

Our study is not without its limitations. Applications with integrated speech recognition and a main server located close by that also utilises artificial intelligence would be more exciting. This would allow the audio input made by the students to be turned to text (speech-to-text), and the most appropriate response from among the responses stored in the database would be pulled up and presented to the student without delay. That said, a large team (software specialist, measurement and evaluation specialist, educational technologist, field specialist, graphic designer, students) would be required to develop such applications, the members of which would need to cooperate to overcome these limitations.

Recommendations to practitioners

Virtual human working groups should be formed and these applications should be integrated into various faculties (nursing, pharmacy, psychology, veterinary) through a cooperation between the health, engineering, and social science fields. The high cost of hiring SPs can be eliminated by incurring a one-time expense for the VP application, and cost-effective solutions may be developed for various scenarios. Because it is very difficult in Turkey to find SPs who can speak English, the scenario and the speech over for the VP application may also be translated to English, helping the large number of foreign students in medical schools develop effective communication.

Recommendations for further studies

VPs that use speech communication and VPs that use written communication can be compared in terms of their effects on student achievement and satisfaction. Scenarios can be translated into different languages and inter-cultural comparisons can be made. VP applications with additional features can be tested, where, for example, the VP also asks questions and the system automatically makes corrections when students fail to ask certain questions. This study examined the logs of the first interviews engaged in by the students, while future studies may examine the development of communication skills over time. Future studies could compare the effectiveness of interviews held via different platforms, such as mobile devices, desktop computers, and tablets. Long-term follow-up studies may be conducted to examine the effects of VPs on student behaviour in real world clinical settings.

References


Powell, E., & Turnbull, J. (2016). Design and implementation of a curriculum to teach communication skills to pediatric critical care fellows. Academic Pediatrics, 16(6), e45-e47. https://doi.org/10.1016/j.acap.2016.05.112


---

Corresponding author: Baris Sezer, barissezer13@hotmail.com