

# The effectiveness of integration of virtual patients in a collaborative learning activity

## Citation for published version (APA):

Marei, H. F., Donkers, J., & Van Merriënboer, J. J. G. (2018). The effectiveness of integration of virtual patients in a collaborative learning activity. *Medical Teacher*, 40, S96-S103.  
<https://doi.org/10.1080/0142159X.2018.1465534>

## Document status and date:

Published: 01/01/2018

## DOI:

[10.1080/0142159X.2018.1465534](https://doi.org/10.1080/0142159X.2018.1465534)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## The effectiveness of integration of virtual patients in a collaborative learning activity

Hesham F. Marei<sup>a</sup> , Jeroen Donkers<sup>b</sup>  and Jeroen J. G. Van Merriënboer<sup>b</sup> 

<sup>a</sup>Department of Biomedical Dental Sciences, College of Dentistry, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia;

<sup>b</sup>Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands

### ABSTRACT

**Background:** Virtual patients (VPs) have been recently integrated within different learning activities.

**Aim:** To compare between the effect of using VPs in a collaborative learning activity and using VPs in an independent learning activity on students' knowledge acquisition, retention and transfer.

**Methods:** For two different topics, respectively 82 and 76 dental students participated in teaching, learning and assessment sessions with VPs. Students from a female campus and from a male campus have been randomly assigned to condition (collaborative and independent), yielding four experimental groups. Each group received a lecture followed by a learning session using two VPs per topic. Students were administered immediate and delayed written tests as well as transfer tests using two VPs to assess their knowledge in diagnosis and treatment.

**Results:** For the treatment items of the immediate and delayed written tests, females outperformed males in the collaborative VP group but not in the independent VP group.

**Conclusion:** On the female campus, the use of VPs in a collaborative learning activity is more effective than its use as an independent learning activity in enhancing students' knowledge acquisition and retention. However, the collaborative use of VPs by itself is not enough to produce consistent results across different groups of students and attention should be given to all the factors that would affect students' interaction.

### Introduction

Virtual patients (VPs) have been defined in the medical literature as a form of computer simulation of real-life clinical scenarios (Ellaway et al. 2006). They can have a linear passive design that is similar to a paper case as used in case-based learning, a linear active design defined as a "string of pearls" that involves a single path through history, examination, diagnosis, and treatment planning, or finally a branching-tree or network design that involves multiple pathways leading to a common end point or multiple end points (Ellaway et al. 2008).

Different studies have shown that feedback to learners, deliberate practice, variability in practice and instructor training are some of the factors that dictate the effectiveness of the use of simulation in general (Cook et al. 2013; Motola et al. 2013). For VPs, it is advocated that the most important impact could come from features that are extraneous to the design of VPs such as how they are aligned with other instructional methods in the curriculum, and whether students should complete VPs in groups or individually. Therefore, there was a call for more research to inform how to effectively use VPs (Cook and Triola 2009). In the same context, Ellaway et al. (2015) indicated that the effectiveness of VPs depends on the activity design that is constructed around them, and the medical teachers should find the best ways to use them.

Recently, VPs have been introduced as an integrated part of a whole learning activity as one VP can be used in many different learning activities, and many VPs can be

### Practice points

- On a female campus, the use of VPs in a collaborative learning activity could enhance students' knowledge acquisition and retention.
- Collaborative learning activities might require a customized VP design that enhances students' interaction.
- The effect of collaborative use of VPs would be more obvious in difficult topics that involve interaction between multiple elements.

used in one activity (Ellaway and Davies 2011). A particular form of integration is called a VP integration scenario (Huwendiek et al. 2013) or VP activity (Ellaway and Davies 2011).

Different VP activity patterns have been presented by Ellaway et al. (2015). Two of which are collaborative learning and independent learning activities. In a collaborative learning activity, learners work in groups with VPs to reach consensus on what their collaborative decisions should be. The role of a tutor in this activity is to control the session, guide the discussion around the VPs and provide feedback to the learners. In an independent learning activity, the learners work individually with VPs, and the responsibility of the tutor is limited to reviewing learner performance at the end of the activity because the correct answers to questions, and the feedback in term of points and scores are already encoded in the software.

Using VPs individually has shown a positive impact on students' knowledge retention when compared with didactic lectures (Botezatu et al. 2010b), but working with VPs in a group might have more advantages because learners share their perspectives, observations and previous experiences, which would allow knowledge to be analyzed, shaped and constructed into new structure.

Multiple studies have confirmed that knowledge acquisition and retention are facilitated by linking new knowledge to prior knowledge to construct meaningful knowledge in a process of elaboration (Dror et al. 2011; Schmidt et al. 2011), and it is well-known that the level of retention of meaningful knowledge is higher than that of meaningless knowledge (Schmidt and Rikers 2007; Schmidt et al. 2011). Moreover, presenting materials in the right context, and even better, in different contexts can lead to improved retention (Custers 2010; Schmidt et al. 2011).

Linking new knowledge to prior knowledge is explained in the theoretical framework of *constructivism*, which attributes learning to the interaction between new knowledge and already existing schemas in long-term memory for the construction of new schemas in a process known as elaboration (Lau 2014; Lunenberg 1998; Tynjala 1999). *Social constructivism* adds the dimension of students' collaboration with each other, and interaction with their tutors as important factors for promoting elaboration (Philpott and Batty 2009). Interacting with others helps to activate students' already existing knowledge and makes it available in short-term memory in order to connect new information from long-term memory to it (Kalyuga 2009).

When students collaborate to apply their knowledge on patient problems, there will be a clustering of knowledge elements that often occur together such as specific history, signs and symptoms of a disease into cognitive schemas in a process of knowledge encapsulation, and involvement of other elements of knowledge such as treatment planning and its outcome will lead to the construction of a new knowledge structure that is known as an "illness script". Therefore, the growth of expertise would depend on the number and varieties of patients' problems or in another sense the number and varieties of constructed illness scripts (Schmidt and Rikers 2007).

Based on this theoretical background, VPs could be seen as an educational tool that can facilitate schema construction, because it provides an authentic context for knowledge application that would allow for elaboration and re-clustering of knowledge relevant to diagnosis, treatment planning, and its outcome into an illness script. Moreover, VPs, especially in collaborative activities would also offer an opportunity for students to work, learn and discuss dilemmas in groups toward a common goal, for instance, completing the VPs correctly.

Although constructing meaningful forms of knowledge can enhance knowledge retention, still the ultimate goal is enabling learners to transfer such knowledge to new situations, and to apply it on problems in the clinical setting. Knowledge transfer involves *near transfer* in which knowledge transfer occurs between different problems in the same context or two closely related contexts, and *far transfer*, that refers to the ability to solve new problems in unfamiliar contexts (Perkins and Salomon 1992; Van Merriënboer 2013).

Research on knowledge transfer advocated that teaching for transfer requires curricula that are rich in authentic activities and full of a variety of problems (Gentile 2000), which is an opportunity that can be offered by VPs (Ellaway et al. 2008). In a study by Botezatu et al. (2010a), undergraduate medical students perceived VPs as an effective tool for both *near transfer* to other types of exams and *far transfer* to real patient settings.

Such prospect that can be obtainable by VPs for knowledge transfer and the current move from using VPs as an isolated e-learning artifact to a more holistic approach that focuses on the pedagogical aspects of VPs as a part of a whole activity triggered us to investigate the effectiveness of using virtual patients in a collaborative learning activity, when compared with their use as independent learning activity, with respect to students' knowledge acquisition, retention, and transfer.

## Method

### Participants

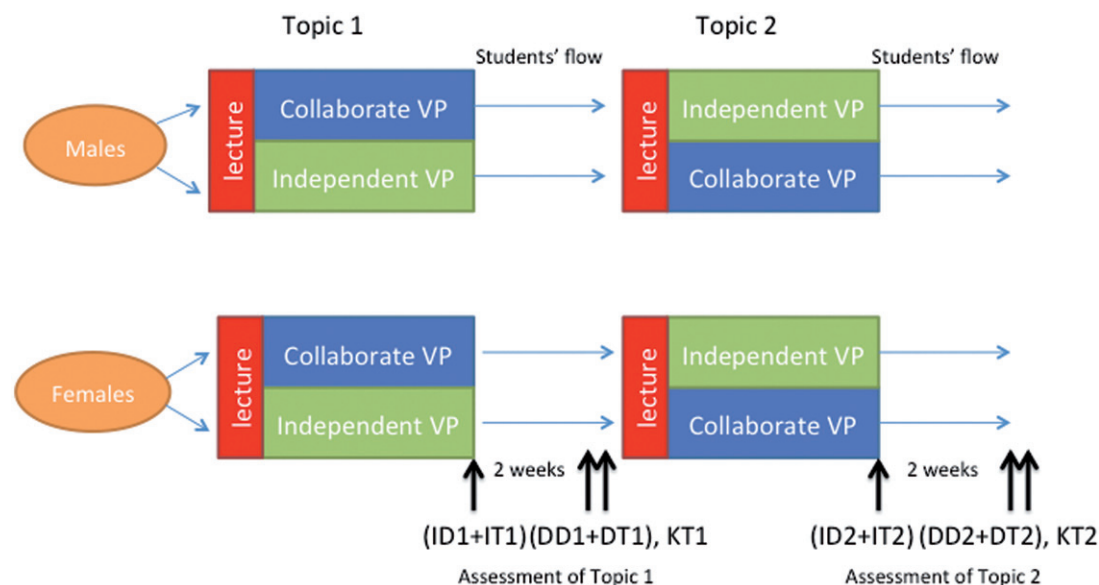
The study was conducted at a dental school that involves two separate campuses for male students and female students. All students who were registered for the first time in the oral and maxillofacial surgery course for the academic year 2015/2016 (1st semester) were invited to participate voluntarily in the study ( $n=96$ ). The study involved the use of eight VPs for teaching, learning and assessment purposes. It targeted two topics, which are impacted wisdom teeth (Topic 1) and maxillofacial trauma (Topic 2) within the same course.

All the dental students were at their 4th academic year of a 6-year discipline-based dental program. Using a stratified random sampling method, male students were randomly distributed using their scores in the previous year as strata to have two comparable groups: male/collaborative group ( $n=23$ ) and male/independent group ( $n=22$ ); female students were also distributed randomly exactly in the same way as male students over two comparable groups: female/collaborative group ( $n=26$ ), and female/independent group ( $n=25$ ). Upon implementation of the study, 14 students have not attended one or more sessions of Topic 1. Therefore, 82 out of 96 students have participated in Topic 1 as follows: male/collaborative group ( $n=18$ ), male/independent group ( $n=17$ ), female/collaborative group ( $n=25$ ) and female/independent group ( $n=22$ ).

For Topic 2, 20 students have not attended one or more sessions. Therefore, the total numbers of students were 76 out of 96 as follows: male/collaborative group ( $n=20$ ), male/independent group ( $n=15$ ), female/collaborative group ( $n=21$ ) and female/independent group ( $n=20$ ).

### Procedure

The interventions consisted of two parts: lecture sessions, which were the same for all four experimental groups, and VP activities, which followed the lecture sessions and were different for the collaborative and independent groups. There were group activities for the male/collaborative and female/collaborative groups, and individual activities for the



**Figure 1.** Flowchart showing the sequence of different interventions. ID1: immediate diagnosis 1; IT1: immediate treatment 1; DD1: delayed diagnosis 1; DT1: delayed treatment 1; KT1: knowledge transfer 1; ID2: immediate diagnosis 2; IT2: immediate treatment 2; DD2: delayed diagnosis 2; DT2: delayed treatment 2; KT2: knowledge transfer 2.

male/independent and female/independent groups (Figure 1).

**Lecture session:** All students received the same lecture that involved a PowerPoint presentation for 90 min that was interrupted by a 10-min break. The lecture followed a teacher-centered approach, and targeted the learning objectives of both topics. They were mainly related to diagnosis of specific conditions through identifying the relevant history, signs and symptoms, required investigations. The lecture also covered formulation of treatment plans through making the students aware of the indications and contraindications of every treatment option, and factors that would affect the outcome of specific surgical procedures. Two subject-matter experts revised the lecture content and structure, and the lecture was delivered by a faculty member (first author) who has been delivering this lecture for the last 3 years to students similar to the participants in this study. After the lecture, the students had another 10-min break, and then commenced the VP session based on their allocation to the experimental groups.

**Collaborative VP learning activity session:** The students were asked to work in groups of five students to reach consensus on what their collaborative decision should be for every question posed by the VPs, which were projected on the classroom screen using a data show projector. The role of the tutor (first author) was to facilitate the session by asking why specific decisions were taken, what the consequence would be of choosing other options, providing feedback and finally operating the VP based on the decisions taken. The VP session lasted for 1 h.

**Independent VP learning activity session:** The same VPs were used as in the collaborative session, but now as an independent learning activity. Each student accessed his/her own computer and worked individually on the VPs from the beginning till the end of the session, which lasted for 1 h. The role of the tutor was limited to monitoring the students during their interaction with the VPs.

Since the lecture session for Topic 1 is identical in both interventions, all male students (independent and collaborative groups) have received the lecture in one lecture

room at the same time, given by the same faculty member (first author), and then they practiced VPs in two classrooms based on the group they were assigned to, then they gathered again in the same lecture room for the immediate tests. The same procedure has been applied to female students, with the same faculty member (first author).

One month later, Topic 2 was presented in the same manner like the first topic. Student groups who have been involved in VPs collaborative activity session were asked to practice VPs independently and vice versa.

## Materials

### Virtual patients

For the teaching and learning sessions, four VPs (two VPs per topic) were used. They had a branched-tree dynamic design and were based on real patient scenarios. Real patient radiographs, laboratory results, intra-oral photographs, and records for other special investigations were used at different stages of VPs path, while two-dimensional graphics were used to represent different clinical settings, and different characters within the VPs (Figure 2).

During practice, each VP branched based on student's choices on crucial decisions such as definitive diagnosis, selecting a specific surgical procedure, or admitting the "patient" to a hospital. The students selected each decision from a shortlist of options. If correct, the student stayed in the main VP path, while if wrong, the student followed the consequences of the decision taken till the next node, where he/she was directed back to the main stream. Other than decision-making in crucial steps, the VP provided immediate feedback to the students on knowledge questions as either correct or incorrect, and briefly explained why an answer was incorrect, while the overall score, which was calculated based on the total number of correct answers, was displayed after the VP had been completed by students.

All VPs are designed, developed and validated, as a part of a research project on the application of VPs in dental





**Figure 2.** Screen shots a, b, c, d from different VPs. (a) virtual characters (dentist, patient and dental assistant) in a dental office during history taking, (b) the learner should navigate through the intra-oral photos of the VP, (c) the learner should take a decision in response to an inquiry from the VP during diagnosis, (d) the learner should take a decision in regard to the treatment plan.

education, by a team that involved subject matter experts, instructional designer, photographer, programmer and animator. VPs development was a collaborative work between faculty members of the dental school and a private e-learning vendor, while partially funded by the Deanship of Scientific Research within our university.

### Measurement instruments

Five tests were constructed for each topic: an immediate diagnosis test (ID), an immediate treatment test (IT), a delayed diagnosis test (DD), a delayed treatment test (DT), and a knowledge transfer test (KT). All tests were paper-based except for the knowledge transfer tests, which were computer-based tests that involved the use of four new VPs (2 VPs for each topic).

The paper-based tests ID and DD are equivalent in format, difficulty and final scores, and the same is true for IT and DT. The immediate tests for Topic 1 (ID1 and IT1) are presented together in one sheet and delivered immediately after the teaching and learning session of Topic 1. They involved 18 Multiple Choice Questions (MCQs), in which each question involves 4 options. The same was true for the immediate tests ID2 and IT2 (18 MCQs), which were delivered immediately after the teaching and learning session for Topic 2. The unannounced delayed tests (DD1, DT1) and (DD2 and DT2) were delivered 2 weeks after the immediate tests of Topic 1 and 2 respectively. They involved the same number and type of items.

All questions were retrieved from the department item bank that contains questions that have been used before and showed a difficulty index of 0.2–0.8 and a discrimination index larger than 0.1. Two subject matter experts established the content validity of test items through test

blueprint and test revision. They also developed and agreed on the scoring key for all questions. The tests were scored electronically. All paper-based tests were constructed to measure factual knowledge (e.g., the complication associated with specific procedure, the risk of damage to certain vital structure, specific terms that describe certain phenomena, the signs and symptoms of specific conditions), conceptual knowledge (e.g. classify an impacted tooth or a specific fracture, estimate the degree of difficulty of specific surgical procedure based on the presented vignette), and strategic knowledge (e.g. request a specific investigation, reach a diagnosis, choose a specific treatment option, prescribe a specific medication based on the patients' condition in the presented vignette).

A factor analysis was performed on the combined paper-based tests to confirm relevant subscales. Two subscales were recognized and included in the analysis: "diagnosis" and "treatment". After deletion of bad items (those with low item-test correlation), the diagnosis scale for Topic 1 consisted of 5 items (Cronbach's  $\alpha = 0.64$ ); three were included in the immediate test (ID1) and 2 were included in the delayed test (DD1), while the diagnosis scale for Topic 2 consisted of 14 items (Cronbach's  $\alpha = 0.66$ ); seven were included in the immediate test (ID2) and seven were included in the delayed test (DD2).

For the treatment scale, deletion of bad items resulted in 11 items for Topic 1 (Cronbach's  $\alpha = 0.70$ ); six items were included in the immediate test (IT1) and five items were included in the delayed test (DT1), while the treatment scale for Topic 2 consisted of 10 items (Cronbach's  $\alpha = 0.70$ ); six items were included in the immediate test (IT2) and four items were included in the delayed test (DT2).

**Table 1.** Means and standard deviations for all scores in percentages, split out by condition and campus.

|                     | Topic 1                        |           |                              |           |                                |           |                              |           |
|---------------------|--------------------------------|-----------|------------------------------|-----------|--------------------------------|-----------|------------------------------|-----------|
|                     | Female campus                  |           |                              |           | Male campus                    |           |                              |           |
|                     | Collaborative ( <i>n</i> = 25) |           | Independent ( <i>n</i> = 22) |           | Collaborative ( <i>n</i> = 18) |           | Independent ( <i>n</i> = 17) |           |
|                     | <i>M</i>                       | <i>SD</i> | <i>M</i>                     | <i>SD</i> | <i>M</i>                       | <i>SD</i> | <i>M</i>                     | <i>SD</i> |
| Immediate diagnosis | 58.6                           | 37.6      | 62.1                         | 44.0      | 51.8                           | 43.1      | 47.0                         | 45.7      |
| Delayed diagnosis   | 58.0                           | 49.3      | 63.6                         | 46.7      | 47.2                           | 43.6      | 52.9                         | 44.9      |
| Immediate treatment | 74.0                           | 27.2      | 55.3                         | 28.8      | 53.7                           | 39.8      | 65.6                         | 34.5      |
| Delayed treatment   | 72.8                           | 22.2      | 54.5                         | 32.7      | 58.8                           | 30.2      | 54.1                         | 28.9      |
| Transfer 1          | 74.5                           | 8.09      | 76.7                         | 8.5       | 67.3                           | 11.7      | 69.8                         | 8.4       |

|                     | Topic 2                        |           |                              |           |                                |           |                              |           |
|---------------------|--------------------------------|-----------|------------------------------|-----------|--------------------------------|-----------|------------------------------|-----------|
|                     | Female campus                  |           |                              |           | Male campus                    |           |                              |           |
|                     | Collaborative ( <i>n</i> = 21) |           | Independent ( <i>n</i> = 20) |           | Collaborative ( <i>n</i> = 20) |           | Independent ( <i>n</i> = 15) |           |
|                     | <i>M</i>                       | <i>SD</i> | <i>M</i>                     | <i>SD</i> | <i>M</i>                       | <i>SD</i> | <i>M</i>                     | <i>SD</i> |
| Immediate diagnosis | 57.8                           | 20.9      | 61.4                         | 15.4      | 57.8                           | 27.9      | 62.8                         | 11.8      |
| Delayed diagnosis   | 50.3                           | 21.4      | 39.2                         | 24.0      | 40.7                           | 31.5      | 52.3                         | 24.5      |
| Immediate treatment | 82.5                           | 17.8      | 57.5                         | 27.8      | 65.8                           | 19.0      | 61.1                         | 14.9      |
| Delayed treatment   | 83.3                           | 29.9      | 36.2                         | 30.8      | 48.7                           | 23.6      | 55.0                         | 16.9      |
| Transfer 2          | 55.3                           | 11.2      | 59.7                         | 12.4      | 62.8                           | 14.6      | 62.4                         | 11.5      |

The computer-based transfer tests KT1 and KT2 were delivered unannounced immediately after the delayed tests. Each test involved two new VPs. The tests targeted mainly strategic knowledge of the relevant topics. The student had to request the appropriate radiograph, interpret the relevant lab test, choose the correct diagnosis, admit the virtual patient to the hospital, select a specific surgical procedure, consent the patient before surgery and finally recommend a management for a specific complication. The VPs' design was similar to the VPs that were used in the teaching and learning sessions, but they provide only a summary report that shows the final score, which is the percentage of correct items to the total number of items in the test. All students worked individually in the transfer tests.

### Data collection and analysis

Data were collected by the college's statistical unit then coded and analyzed using SPSS Software (SPSS 22.0, SPSS Inc., Chicago, IL). For both Topic 1 and Topic 2,  $2 \times 2$  ANOVAs with the factors campus (male, female) and condition (collaborative, independent) were applied to test for main effects and interactions on immediate diagnosis, delayed diagnosis, immediate treatment, delayed treatment, and transfer scores. A *p* value less than 0.05 is considered significant.

## Results

Eighty-two (47 females and 35 males) out of 96 students have participated in all the teaching, learning, and assessment sessions of Topic 1, while 76 students (41 females and 35 males) have participated in Topic 2 sessions.

Table 1 provides the means and standard deviations for all scores, split out for condition (collaborative, independent) and campus (male, female).

For the immediate and delayed diagnosis scores of Topic 1 and Topic 2, there were no significant effects of campus, condition or their interaction.

For the immediate treatment score for Topic 1, a significant interaction between condition and campus was found,  $F(1, 78) = 4.51, p < 0.05$ . As shown in the left upper graph of Figure 3(a), females ( $M = 74.0, SD = 27.2$ ) do better than

males ( $M = 53.7, SD = 39.8$ ) in the collaborative condition, but females ( $M = 55.3, SD = 28.2$ ) are not superior to males ( $M = 65.6, SD = 34.5$ ) in the independent condition. Thus, females are superior to males *only* in the collaborative condition. The main effect of campus is not significant,  $F(1, 78) = 0.47, p = 0.49$ , and the main effect of the condition is also not significant  $F(1, 78) = 0.21, p = 0.64$ .

A similar pattern was found for the immediate treatment score for Topic 2; again, a significant interaction between condition and campus was found. As shown in the right upper graph of Figure 3(c), females ( $M = 82.5, SD = 17.8$ ) do better than males ( $M = 65.8, SD = 19.0$ ) in the collaborative condition, but females ( $M = 57.5, SD = 27.8$ ) are not superior to males ( $M = 61.1, SD = 14.9$ ) in the independent condition. This interaction effect is significant,  $F(1, 72) = 4.45, p < 0.05$ . The main effect of campus is not significant,  $F(1, 72) = 1.84, p = 0.17$ , while the main effect of condition is significant,  $F(1, 72) = 9.55, p < 0.05$ . However, as can be seen in the right upper graph of Figure 3, this main effect of condition cannot be interpreted due to the disordinal interaction.

For the delayed treatment score of Topic 1, there was a similar trend of results. As shown in the left lower graph of Figure 3(b), females ( $M = 72.8, SD = 22.2$ ) seemed to score somewhat higher than males ( $M = 58.8, SD = 30.2$ ) in the collaborative condition but not in the independent condition, but this interaction was not statistically significant,  $F(1, 78) = 1.11, p = 0.293$ . The main effects of campus ( $p = 0.26$ ) and condition ( $p = 0.07$ ) are also not significant.

For the delayed treatment score of Topic 2, again a significant interaction between campus and condition was found. As shown in the right lower graph of Figure 3(d), females ( $M = 83.3, SD = 29.9$ ) do better than males ( $M = 48.7, SD = 23.6$ ) in the collaborative condition, but they ( $M = 36.2, SD = 30.8$ ) are not superior to males ( $M = 55.0, SD = 16.9$ ) in the independent condition. This interaction effect is significant,  $F(1, 72) = 18.8, p < 0.001$ . The main effect of campus is not significant,  $F(1, 72) = 1.6, p = 0.2$ , while the main effect of condition is significant,  $F(1, 72) = 11.0, p < 0.05$ . But as for the immediate treatment score of Topic 2, the main effect of condition cannot sensibly be interpreted because there is a disordinal interaction.

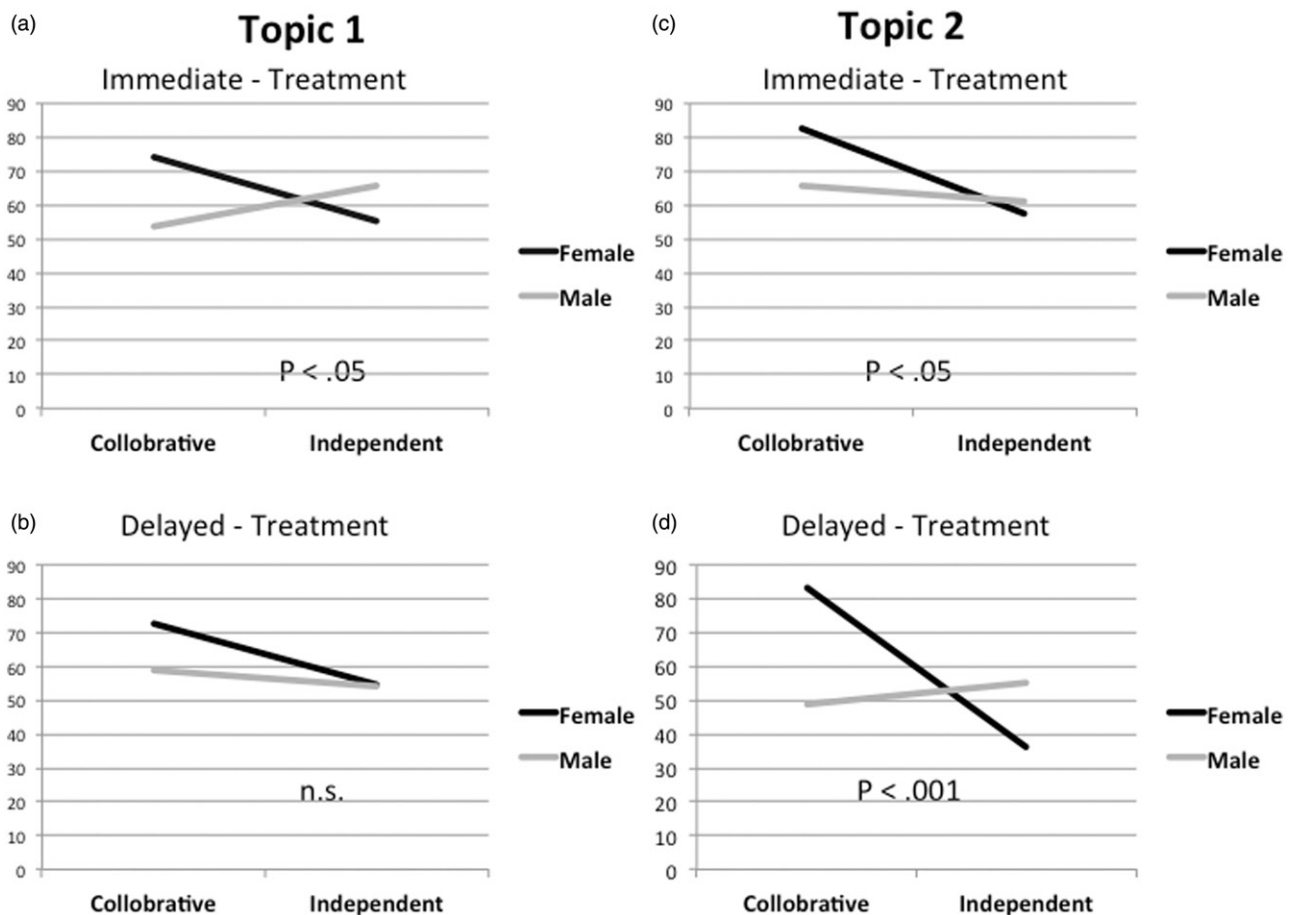


Figure 3. Students' scores in the treatment tests of topic 1 and 2 per condition and campus.

For knowledge transfer, a significant main effect was found for campus on students' scores in Topic 1. Females ( $M=75.5$ ,  $SD=8.31$ ) do better than males ( $M=68.5$ ,  $SD=10.1$ ),  $F(1, 78)=11.5$ ,  $p<0.05$ . There were no significant effects for condition ( $p=0.25$ ) or the interaction between campus and condition on student scores. The same is true for Topic 2 as there were no significant effects of campus, condition or their interaction.

## Discussion

This study compared the effect of collaborative and independent use of VPs on students' knowledge acquisition, retention and transfer across two campuses, namely a male and female campus. Our hypothesis was partially supported, as the impact of collaborative use of VPs was clear in the female campus only and mainly along knowledge acquisition and retention of treatment items.

As shown in Figure 3, females who worked collaboratively with the VPs were better on the treatment tests than females who worked independently, but these results were not attained on the male campus. This finding could possibly be attributed to a higher amount of discussion and interaction occurring between females. It was shown by Adrianson (2001) that females have a tendency to produce more messages in face-to-face communication, and there were also more opinion changes from females than from males in collaborative learning. The same conclusion was drawn in asynchronized online communication by Savicki et al. (1996), as the authors found that female groups were the most satisfied by the group process and changed their

opinions most as a result of group activity. Moreover, it was shown that gender grouping mainly influences students' attitudes toward collaborative learning as female students perform better in single gender grouping, while males perform better in mixed gender grouping (Zhan et al. 2015).

Moreover, the division of students in male and female campuses could have created conditions in the female campus that are more conducive to collaboration and peer interaction. As groups' creation for the collaborative activities was constructed randomly by the students themselves by forming smaller groups of five students each. Therefore, we suspect that the construction of these small groups has created a more favorable mixture in the female campus than in the male campus. In other words, female collaborative small groups might have involved a mixture of high-aptitude students with one or more low-aptitude students. Such mixture allowed high-aptitude students to provide scaffolding to lower-aptitude students. It has been documented that low-aptitude students can cross the zone of proximal development (ZPD) when help and scaffolding are provided in the appropriate time. ZPD is the difference between what a learner can do without help and what he or she can do with help (Vygotsky 1980). de Leng et al. (2009) showed that for achieving effective use of computer-based simulated cases, it is important to introduce scaffolding strategies such as guided peer interaction.

The question that requires an answer now is why the superiority of collaborative performance was evident only in treatment scores and such effect was absent in diagnosis and transfer scores. This could be explained based on the

nature of the content that was provided to the students. In Topics 1 and 2, each topic presented one main dental problem that involved different treatment options, and surgical approaches based on different inputs such as the age of the patients and their underlying medical history. That is to say, it was easy for the students to construct their own diagnosis knowledge structure independently as it was only one problem per topic. But in knowledge that is relevant to treatment, where there were high amounts of interrelating elements, the interaction between students facilitated the elaboration; therefore, the collaborative performance showed its impact.

This study has provided a preliminary conclusion about the effectiveness of collaborative use of VPs on knowledge acquisition and retention. Future studies could investigate the effect of multiple and longer exposure to collaborative VPs activities, for instance at the end of a whole semester. Such context might allow defining and assessing specific metrics of student performance, which could clarify when and how the learner moves from novice to expert.

The study had limitations in terms of the amount of exposure of our students to learning using VPs, as it was only one session for 1 h, but it was necessary to standardize the time of exposure across different groups and eliminate any bias that might occur as a result of studying at home especially before the immediate test. Another important limitation is that the results provide little input to universities with gender mixed groups of students but we suspect that our results might be generalized to educational situations, where there are gender-separated groups of students.

## Conclusions

On the female campus, the use of VPs in a collaborative learning activity is more effective than its use in an independent learning activity in enhancing students' knowledge acquisition and retention related to treatment decisions. However, the collaborative use of VPs by itself is not enough to produce consistent results across different groups of students and attention should be given to all the factors that would affect students' interaction.

## Ethical approval

The study was reviewed and approved by the University Ethical Committee. IRB number: IRB-2015-02-099.

## Acknowledgments

Authors would like to thank Dr. Mohamed AL-Eraky from the Medical Education Department, and all the members of Oral and Maxillofacial Surgery Division who revised the lectures content and structure and established the content validity of the different tests. The Saudi Supplement on Medical Education is sponsored and supported by Al Imam Mohammed Ibn Saud Islamic University.

## Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

## Notes on contributors


**Hesham F. Marei**, PhD, is an Associate Professor and Consultant of Oral and Maxillofacial Surgery at Imam Abdulrahman Bin Faisal University, Saudi Arabia. Dr. Marei is currently a PhD fellow at Maastricht University, Netherlands.


**Jeroen Donkers**, PhD, is an Assistant Professor, at the Department of Educational Development and Research, Faculty of Health, Medicine, and Life Sciences, Maastricht University, Netherlands.

**Jeroen J. G. Van Merrienboer**, PhD, is a Professor of Learning and Instruction at the Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University. He is the Program Director of Research in Education, School of Health Professions Education and Scientific Director of the Interuniversity Center for Educational Research.

## ORCID

Hesham F. Marei  <http://orcid.org/0000-0002-5967-6473>

Jeroen Donkers  <http://orcid.org/0000-0002-6769-0355>

Jeroen J. G. Van Merrienboer  <http://orcid.org/0000-0002-5868-7031>

## References

- Adrianson L. 2001. Gender and computer-mediated communication: Group processes in problem solving. *Comput Human Behav.* 17:71–94.
- Botezatu M, Hult H, Fors UG. 2010a. Virtual patient simulation: what do students make of it? A focus group study. *BMC Med Educ.* 10:91.
- Botezatu M, Hult H, Tessma MK, Fors U. 2010b. Virtual patient simulation: knowledge gain or knowledge loss? *Med Teach.* 32:562–568.
- Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hatala R. 2013. Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Med Teach.* 35:e867–e898.
- Cook DA, Triola MM. 2009. Virtual patients: a critical literature review and proposed next steps. *Med Educ.* 43:303–311.
- Custers EJ. 2010. Long-term retention of basic science knowledge: a review study. *Adv in Health Sci Educ.* 15:109–128.
- de Leng BA, Muijtens AM, van der Vleuten CP. 2009. The effect of face-to-face collaborative learning on the elaboration of computer-based simulated cases. *Simul Healthc.* 4:217–222.
- Dror I, Schmidt P, O'Connor L. 2011. A cognitive perspective on technology enhanced learning in medical training: great opportunities, pitfalls and challenges. *Med Teach.* 33:291–296.
- Ellaway R, Davies D. 2011. Design for learning: Deconstructing virtual patient activities. *Med Teach.* 33:303–310.
- Ellaway R, Poulton T, Fors U, McGee JB, Albright S. 2008. Building a virtual patient commons. *Med Teach.* 30:170–174.
- Ellaway R, Topps D, Lee S, Armson H. 2015. Virtual patient activity patterns for clinical learning. *Clin Teach.* 12:267–271.
- Ellaway R, Candler C, Greene P, Smothers V. 2006. An architectural model for medbiquitous virtual patients. Baltimore, MD: MedBiquitous.
- Gentile JR. 2000. Learning, transfer of. In: Kazdin AE, Kazdin AE, editors. *Encyclopedia of psychology*, vol 5. Washington, DC; New York, NY: American Psychological Association, Oxford University Press. p. 13–16.
- Huwendiek S, Duncker C, Reichert F, De Leng BA, Dolmans D, van der Vleuten CP, Haag M, Hoffmann GF, Tonshoff B. 2013. Learner preferences regarding integrating, sequencing and aligning virtual patients with other activities in the undergraduate medical curriculum: A focus group study. *Med Teach.* 35:920–929.
- Kalyuga S. 2009. Knowledge elaboration: a cognitive load perspective. *Learn Instruct.* 19:402–410.
- Lau KH. 2014. Computer-based teaching module design: principles derived from learning theories. *Med Educ.* 48:247–254.
- Lunenburg FC. 1998. Constructivism and technology: instructional designs for successful education reform. *J Instruct Psychol.* 25:75–82.
- Motola I, Devine LA, Chung HS, Sullivan JE, Issenberg SB. 2013. Simulation in healthcare education: a best evidence practical guide. *Amee guide no. 82. Med Teach.* 35:e1511–e1530.
- Perkins DN, Salomon G. 1992. Transfer of learning. *Int Encyclopedia Educ.* 2:6452–6457.



- Philpott J, Batty H. 2009. Learning best together: social constructivism and global partnerships in medical education. *Med Educ.* 43:923–924.
- Savicki V, Kelley M, Lingenfelter D. 1996. Gender and group composition in small task groups using computer-mediated communication. *Comput Human Behav.* 12:209–224.
- Schmidt HG, Rikers RM. 2007. How expertise develops in medicine: Knowledge encapsulation and illness script formation. *Med Educ.* 41:1133–1139.
- Schmidt HG, Rotgans JI, Yew EH. 2011. The process of problem-based learning: What works and why. *Med Educ.* 45:792–806.
- Tynjala P. 1999. Towards expert knowledge? A comparison between a constructivist and a traditional learning environment in the university. *Int J Educ Res.* 31:357–442.
- Van Merriënboer JJ. 2013. Perspectives on problem solving and instruction. *Comput Educ.* 64:153–160.
- Vygotsky LS. 1980. *Mind in society: the development of higher psychological processes.* Cambridge, MA: Harvard university press.
- Zhan Z, Fong PS, Mei H, Liang T. 2015. Effects of gender grouping on students' group performance, individual achievements and attitudes in computer-supported collaborative learning. *Comput Human Behav.* 48:587–596.