




Collaborative use of virtual patients after a lecture enhances learning with minimal investment of cognitive load

Hesham F. Marei, Jeroen Donkers, Mohamed M. Al-Eraky & Jeroen J. G. Van Merriënboer

To cite this article: Hesham F. Marei, Jeroen Donkers, Mohamed M. Al-Eraky & Jeroen J. G. Van Merriënboer (2019) Collaborative use of virtual patients after a lecture enhances learning with minimal investment of cognitive load, *Medical Teacher*, 41:3, 332-339, DOI: [10.1080/0142159X.2018.1472372](https://doi.org/10.1080/0142159X.2018.1472372)

To link to this article: <https://doi.org/10.1080/0142159X.2018.1472372>




View supplementary material 



Published online: 25 May 2018.



Submit your article to this journal 



Article views: 656



View related articles 



View Crossmark data 



Citing articles: 7 View citing articles 



Collaborative use of virtual patients after a lecture enhances learning with minimal investment of cognitive load

Hesham F. Marei^a , Jeroen Donkers^b , Mohamed M. Al-Eraky^c and Jeroen J. G. Van Merriënboer^b

^aDepartment of Biomedical Dental Sciences, College of Dentistry, Imam Abdulrahman Bin Faisal University, Saudi Arabia; ^bSchool of Health Professions Education, Maastricht University, the Netherlands; ^cDevelopment and Academic Initiatives Directorate, Imam Abdulrahman Bin Faisal University, Saudi Arabia

ABSTRACT

Background: The use of virtual patients (VPs), due to their high complexity and/or inappropriate sequencing with other instructional methods, might cause a high cognitive load, which hampers learning.

Aim: To investigate the efficiency of instructional methods that involved three different applications of VPs combined with lectures.

Method: From two consecutive batches, 171 out of 183 students have participated in lecture and VPs sessions. One group received a lecture session followed by a collaborative VPs learning activity (collaborative deductive). The other two groups received a lecture session and an independent VP learning activity, which either followed the lecture session (independent deductive) or preceded it (independent inductive). All groups were administered written knowledge acquisition and retention tests as well as transfer tests using two new VPs. All participants completed a cognitive load questionnaire, which measured intrinsic, extraneous and germane load. Mixed effect analysis of cognitive load and efficiency using the R statistical program was performed.

Results: The highest intrinsic and extraneous load was found in the independent inductive group, while the lowest intrinsic and extraneous load was seen in the collaborative deductive group. Furthermore, comparisons showed a significantly higher efficiency, that is, higher performance in combination with lower cognitive load, for the collaborative deductive group than for the other two groups.

Conclusion: Collaborative use of VPs after a lecture is the most efficient instructional method, of those tested, as it leads to better learning and transfer combined with lower cognitive load, when compared with independent use of VPs, either before or after the lecture.

Introduction

A virtual patient (VP) is a computer simulation of a real-life clinical scenario for the purpose of teaching, learning and assessment (Ellaway et al. 2006). Originally, VPs were seen as a standalone e-learning tool that can be accessed by students via an e-learning management system (Berman et al. 2009). Learning from VPs, as a standalone activity, depends on the inherent characteristics that are embedded in their design, and on the time devoted by students to use them (Huwendiek et al. 2009; Bateman et al. 2013). Recently, more emphasis has been given to the value of VPs when they are integrated and sequenced with other instructional methods such as collaborative learning or lecturing (Ellaway and Davies 2011; Edelbring et al. 2012; Ellaway et al. 2015).

The use of VPs, due to their high complexity or element interactivity, might cause a high intrinsic cognitive load. In case of cognitive overload due to the complexity of the learning materials, learning is impaired (Van Merriënboer and Sweller 2010). Moreover, the sequence of complex VPs with other instructional methods that do not respect the level of prior knowledge of the learners could result in poor instructions, which is another potential source of cognitive load known as extraneous cognitive load (Kirschner et al. 2006).

Ideally, VPs sessions should be designed in such a way that cognitive overload is prevented. This can be reached

Practice points

- The relevant theory should be presented first and then followed by a collaborative use of VPs.
- Future selection of instructional methods involving VPs should consider a cognitive load measurement, as some VPs applications might be detrimental to learning.
- The amount of guidance offered during VPs practice should be based on the complexity of the provided problem and the experience of the learners.
- The intended use of VPs should be incorporated and considered during the initial stages of VPs development.

by reducing extraneous load (which is always desirable) and/or intrinsic load (which might be necessary when learning tasks are too complex given the prior knowledge of the learners). By preventing cognitive overload, germane processing is facilitated, meaning that learners can optimally allocate their available working memory resources to deal with learning processes such as schema construction and schema automation (Van Merriënboer and Sweller 2010).

The extraneous cognitive load refers to the cognitive load that is not necessary for learning and/or execution of the task (Young et al. 2014). PowerPoint presentations that use slides full of texts and written descriptions of images is an example of poor instruction that causes high extraneous load (Van Merriënboer and Sweller 2010). On the other side, intrinsic cognitive load is the load necessary for executing the learning tasks and it increases with the increase in the number of information elements and the extent to which these elements interact with each other (Young et al. 2014). Novice students approaching a surgical problem for a patient with multiple interacting systematic diseases before facing it on a medically fit patient is an example of imposing (too) high intrinsic cognitive load on students. Therefore, it is important to adapt the complexity of the problem to the prior knowledge of the learners by sequencing the problems from simple to complex (Van Merriënboer and Sweller 2010).

To reduce intrinsic cognitive load, the four components instructional design model (4C/ID; Van Merriënboer and Kirschner 2018) suggests that supportive information (theoretical information that is relevant to performing the learning task) should be tailored and provided before executing specific learning tasks for learners, who already have some relevant prior knowledge in the domain. This deductive approach helps the learner to construct the cognitive schemata that are necessary to carry out problem solving, reasoning and decision making while completing the learning tasks (Vandewaetere et al. 2015).

The sequence of VPs (learning task) and lecture (supportive information) has been evaluated with respect to students' preference and effectiveness (Huwendiek et al. 2013; Marei et al. 2017). Huwendiek et al. (2013) showed that fifth-year students preferred having a lecture that is followed by VPs, a small group discussion, and then a real patient session. Marei et al. (2017) reported no significant difference between sequencing VPs before the lecture (inductive learning approach) or after it (deductive learning approach) on students' performance when direct instruction is offered in the inductive approach to novice learners.

Collaborative learning represents one form of scaffolding that is offered by group members to each other during the execution of learning tasks. Kirschner et al. (2009a) argued that assigning high complexity tasks to groups of learners allows information to be divided across a larger reservoir of cognitive capacity and might result in more effective and efficient learning than assigning them to an individual learner. However, if the task is simple, learning might be impaired as a result of group members' interaction as the discussion might impose an extraneous load that is not necessary for execution of a task that can be accomplished easily by independent work (Kirschner et al. 2009a).

The current development of instruments that can be used at the end of teaching and learning sessions to measure the different types of cognitive load (Brunken et al. 2003; Leppink et al. 2013) has made the practice of exposing the students to cognitive overload educationally unjustified. Cognitive load measurements could be even of higher priority when new instructional methods like VPs are implemented or whenever there is a curriculum change that involves a new learning content. Therefore, it is important to investigate the most efficient way to integrate VPs in educational practice. Such practice would be

most effective at managing learners' cognitive load and at the same time has the maximum effect on learner's performance.

Van Gog and Paas (2008) advocated that drawing conclusions about the efficiency of instructional methods emerges not only from test performance but also from cognitive load measurements during learning. In the same context, Leppink (2016) stated that interpretation of cognitive load measures in isolation from learning outcomes has no meaning in advocating specific instructional methods because lower cognitive load is not a goal in itself, but it should be reflected in better learning thus higher learning outcomes. Multiple studies used *instructional efficiency measures* to examine instructional efficiency in terms of the learning process, that is, combining the cognitive load investment during instructions with the test performance attained as a result of these instructions (Van Gog and Paas 2008).

The purpose of this study is to investigate the efficiency of instructional methods that involved three different applications of VPs combined with lectures. Our hypothesis is that collaborative use of VPs after the lecture (i.e. deductive) would be the most efficient instructional method when compared with independent use of VPs either before (i.e. inductive) or after the lecture.

Method

The study was conducted in a dental school over two consecutive academic years. The dental school involves two separate campuses for male students and female students. However, in the current study, all the teaching and learning sessions were conducted in the same classroom but at different times for the male and female students. The teaching and learning sessions targeted two topics, which are impacted wisdom teeth (Topic 1) and maxillofacial trauma (Topic 2) within the same course.

Participants

All students who were registered in the oral and maxillofacial surgery course for the academic years commenced on 2015 (year 1) and 2016 (year 2) were invited to participate voluntarily in the study ($N=183$). The students were randomly assigned in year 1 into two different interventions, which are collaborative deductive ($n=49$) and independent deductive year 1 ($n=47$), while in year 2; they were randomly assigned into independent deductive ($n=43$) and independent inductive ($n=44$). Thus, for both years combined, $n=49$ for the collaborative deductive group, $n=90$ for the independent deductive group, and $n=44$ for the independent inductive group. The students are familiar with problem based learning as an instructional method. They practiced it along different courses in the curriculum. However, the most dominant instructional method in our curriculum is lectures.

Materials

Teaching and learning sessions

The teaching and learning sessions consisted of two parts: a lecture session, which was the same per topic for all

groups, and a VP learning session which lasted for one hour and differs per group as follows:

Collaborative Deductive (CD) VP session: After the lecture, the students of year 1 were asked to work in groups of five students (single-gender groups) to discuss and take a collective decision for every question posed by the VPs, which were projected on the classroom screen using a data show projector. The role of the tutor was to facilitate the session by asking why specific decisions were taken, providing feedback and finally operating the VP based on the decisions taken.

Independent Deductive (ID) VP session: After the lecture, each student of year 1 and 2 accessed the university E-learning management system and worked individually on the VPs. The role of the tutor was limited to monitor the students during their interaction with the VPs.

Independent Inductive (II) VP session: Before the lecture, each student of year 2 accessed the university E-learning management system and worked individually on the VPs in the same manner as in the independent deductive group.

Lecture session: The same lecture was presented to all students. It was in the form of a PowerPoint presentation for 90 minutes with a 10-minutes break halfway. The lecture was designed to target diagnosis and treatment of different conditions including management of complications that might arise as a result of different surgical procedures. The lecture content and structure was revised by two subject-matter experts and was delivered by an expert faculty member to both the male and female groups.

Virtual patients

Four VPs (two VPs per topic) were used in the VP learning activity sessions. The VPs had a branched-tree dynamic design, and were based on real patient scenarios. Real patient radiographs, lab results, intra-oral photos, 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 ([supplemental Appendix 1](#)).

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 (in ID and II) or students (in CD) stayed in the main VP path, while if wrong, the student/s followed the consequences of the decision taken till the next node and were then directed back to the main stream. The VP provided immediate feedback to the students on knowledge questions as either correct or 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.

Questionnaire

The study applied the cognitive load questionnaire ([Table 1](#)) developed by Leppink et al. (2014). The questionnaire was applied in multiple previous studies to measure intrinsic, extraneous, and germane cognitive load (Tremblay et al. 2017). The questionnaire involves 13 items with 11-point (0–10) rating scales. The first 4 items measure intrinsic cognitive load (IL), items 5 to 8 measure extraneous cognitive

load (EL), while the last 5 items measure germane load (GL). The items were adapted to our study by changing the word “activity” in the original questionnaire into “virtual patient activity” to refer specifically to the use of VPs. The scores for the three types of load were scaled so that each type of load ranged between 0 and 1.

Knowledge tests

Three tests were delivered for each topic: A Knowledge Acquisition test (KA), a Knowledge Retention test (KR), and a Knowledge Transfer test (KT). KA and KR were paper-based tests that are similar in format, difficulty and number of items, while KT was a computer-based test that involved the use of four VPs (2 VPs for each of the two topics). Each test consisted of 18 to 25 Multiple Choice Questions (MCQs), in which each question involved four answer options. For the KA and KR tests, we used the same blueprint and difficulty selection. The items were retrieved from an item bank that involves items that were tested and used before. Furthermore, the tests as a whole were piloted on a group of similar students before using them in the study. Then, the items are subjected to improvement and revision (content validation) by a group of subject matter experts ([supplemental Appendix 2](#)). The internal consistency (Cronbach's alpha) of items of Topic 1 and 2 over the two years was higher than 0.7.

The computer-based transfer tests KT involved two new VPs for each topic. Two independent subject matter experts validated the new VPs to ensure that they presented new situations that are relevant to the VPs used in the learning sessions. The students have independently taken the transfer test by logging in to their account on the university learning management system. The VP software did all scoring automatically.

Procedure

For Topic 1, the study involved the implementation of the lecture and VPs session as per the assigned groups. One month later, Topic 2 was presented in the same manner as the first topic. Student groups of year 1 that have been involved in CD session were asked to practice VPs independently and vice versa. The same applied to student groups of year 2, as those who have been involved in ID session in Topic 1 were now asked to practice the VPs before the lecture session in Topic 2 and vice versa.

Immediately after each VP session, participants were asked to complete the cognitive load questionnaire over 20 minutes. The questionnaire was delivered in its own original English version. The students handed the questionnaire back immediately after completing it. Confirmatory factor analysis (CFA) was applied to check to what extent the pre-validated questionnaire fits to the current context. It showed the following measures: SRMR = 0.058, RMSEA = 0.071, CFI = 0.980, TLI = 0.974, which indicates a reasonable fit.

The KA tests for Topics 1 and 2 were delivered immediately after the lecture and VP sessions of Topic 1 and 2, respectively. The KR tests were unannounced and delivered two weeks after the immediate KA tests of Topic 1 and 2, respectively. The computer-based transfer tests KT for Topic 1 and 2 were delivered unannounced immediately after the retention tests. The results of the tests were displayed on

Table 1. Cognitive load questionnaire by Leppink et al. (2014).

1. The content of this VP activity was very complex.
2. The problem/s covered in this VP activity was/were very complex.
3. In this VP activity, very complex terms were mentioned.
4. I invested a very high mental effort in the complexity of this VP activity.
5. The explanations and instructions in this VP activity were very unclear.
6. The explanations and instructions in this VP activity were full of unclear language.
7. The explanations and instructions in this VP activity were, in terms of learning, very ineffective.
8. I invested a very high mental effort in unclear and ineffective explanations and instructions in this VP activity.
9. This VP activity really enhanced my understanding of the content that was covered.
10. This VP activity really enhanced my understanding of the problem/s that was/were covered.
11. This VP activity really enhanced my knowledge of the terms that were mentioned.
12. This VP activity really enhanced my knowledge and understanding of how to deal with the problem/s covered.
13. I invested a very high mental effort during this VP activity in enhancing my knowledge and understanding.

Please take your time to read each of the questions carefully and respond to each of them on a scale from 0 to 10, in which "0" means not at all and "10" means completely the case.

the last screen at the end of the VP and were collected by the tutor.

Data collection and analysis

The college's statistical unit collected and tabulated the data in an Excel sheet. An expert biostatistician (second author) analyzed the data with the R statistical program (R Core Team 2017).

Mixed effect analysis of cognitive load and efficiency: In the linear mixed effect analyses using R (R Core Team 2017) and lme4 (Bates et al. 2014), group (treatment) and gender were taken as fixed effects, and the intercepts of topic (per study) and student as random effects. Interactions between gender and group were only included when significant. *p* values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. Normality and homoscedasticity were checked for each model. AIC, conditional R^2 and marginal R^2 were obtained as measures of fit.

Efficiency (*E*) is computed by first computing z-scores of cognitive load (IL and EL) and test scores (KA, KR, KT) per study x topic (so in four parts). Z-scores are computed by taking the mean and SD and computing $(x - \text{mean})/\text{SD}$ for each *x*. The efficiency of each group is then computed by taking $(z_P - z_L)/\sqrt{2}$, where z_P is the z-score of the performance (test score) and z_L is the z-score of the cognitive load. In a system of axes, the efficiency of each group is thus computed as the shortest distance from a point in space to the line $z_P = z_L$: if performance is higher than can be expected on the basis of cognitive load, efficiency is positive; if performance is lower than can be expected on the basis of cognitive load, efficiency is negative (for a full description of the computations, see Paas and Van Merriënboer 1993).

This resulted in 6 efficiency measures for the three different instructions (CD, ID, II) as follows: Efficiency intrinsic load & KA (EILKA) = $z_{KA} - z_{IL}/\sqrt{2}$. The same was applied to calculate the other efficiency measures: EELKA (extraneous load & KA), EILKR (intrinsic load & KR), EELKR (extraneous load & KR), EILKT (intrinsic load & KT), and EELKT (extraneous load & KT).

Results

The study involved the participation of 171 out of 183 invited students (91 females, 80 males) yielding 320 data records (173 females, 147 males) as follow: 80 records in group CD, 156 records in group ID and 84 records in group

II. The data record is the students' scores of KA, KR and KT tests of Topic 1 and/or Topic 2 in addition to a completely filled cognitive load questionnaire. There was a dropout of 22 data records due to incomplete data. There was no significant effect of the gender on students' scores and cognitive load scores.

The study showed that the highest IL and EL was found in the II group, while the lowest IL and EL was seen in the CD group (Figure 1). In regard to germane load, the highest was in the CD group, while the lowest was found in the II group. There was also a high correlation between IL and EL (0.704), and medium negative correlations between IL and GL (−0.491). The same was true for EL and GL (−0.422). The correlations between IL, EL and KA, KR scores are close to zero. IL correlates medium with KT (0.45), while EL has a lower correlation with KT (0.31). GL has medium correlation with KA (0.56) and KT (0.45) and lower correlation with KR (0.22).

The results of the six efficiency measures: EILKA, EELKA, EILKR, EELKR, EILKT, and EELKT are presented per group and gender in Figure 2. There is a general trend of higher scores for the CD group than for the ID and II groups.

Mixed effect analysis showed a significant change in IL, EL and GL when the CD learning approach is compared with both ID and II approaches. Comparing CD to ID resulted in a significantly higher IL (0.116 ± 0.026) and EL (0.147 ± 0.026), while a significantly lower GL (-0.086 ± 0.021) in the ID group. The same was true for comparing CD to II, as there was a significantly higher IL (0.262 ± 0.036) and EL (0.230 ± 0.035), while a significantly lower GL (-0.202 ± 0.029) in the II group (Table 2). The model fits for IL, EL and GL were reasonable (AICs were −188.2, −211.1 and −377.1, respectively, Marginal R^2 s were 0.14, 0.14 and 0.15, conditional R^2 s were 0.63, 0.48, 0.56, respectively).

There was a significant change in the efficiency of instruction when CD learning approach is compared with both ID and II approaches. Comparison of ID and CD showed a significantly higher efficiency for the CD group. The same effect was found when CD is compared to II (Table 3). The model fits for the efficiency measures were, however, not strong (AICs ranged from 881 to 929, marginal R^2 s between 0.08 and 0.14, but conditional R^2 s varied. For EELKA and EILKA, conditional R^2 was 0.44 and 0.49, for the other measures, the conditional R^2 ranged between 0.20 and 0.30).

Discussion

The current study investigated the efficiency of instructional methods that involved three different applications of

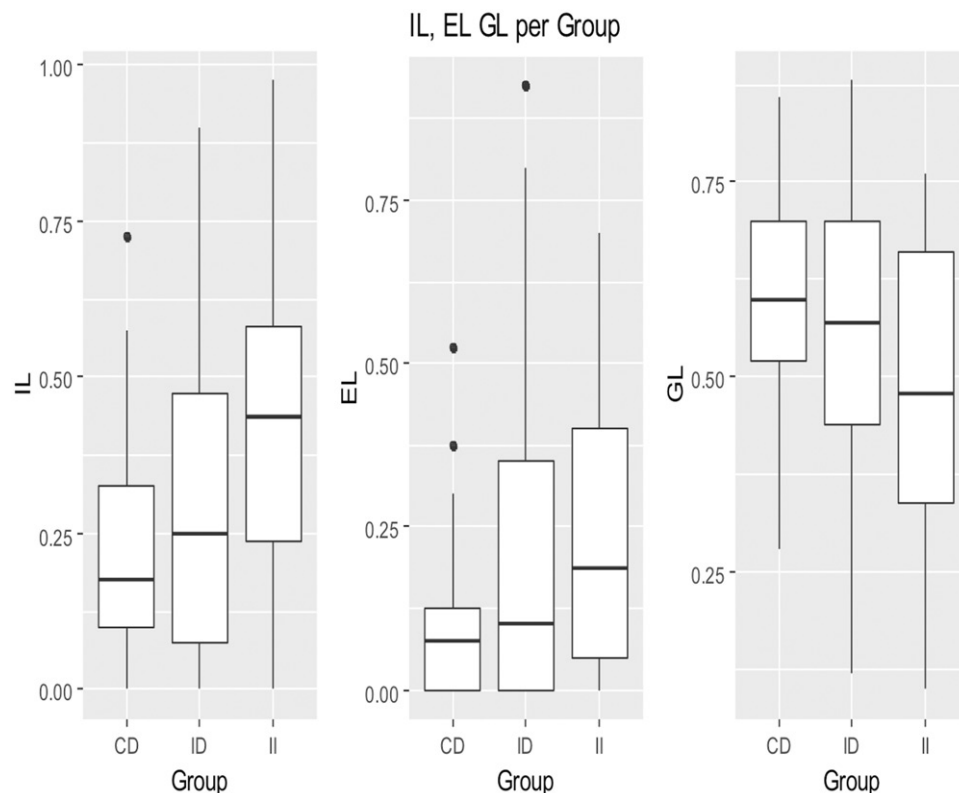


Figure 1. The invested cognitive load per group. IL: intrinsic load; EL: extraneous load; GL: germane load; CD: collaborative deductive; ID: independent deductive and II: independent inductive. The central horizontal line represents the median; the top and bottom of the boxes represent the 25th and 75th percentile. The lines on top and below the boxes extend no further than 1.5 times the interquartile range outside the box. Data points further away are considered as outliers and are indicated as dots.

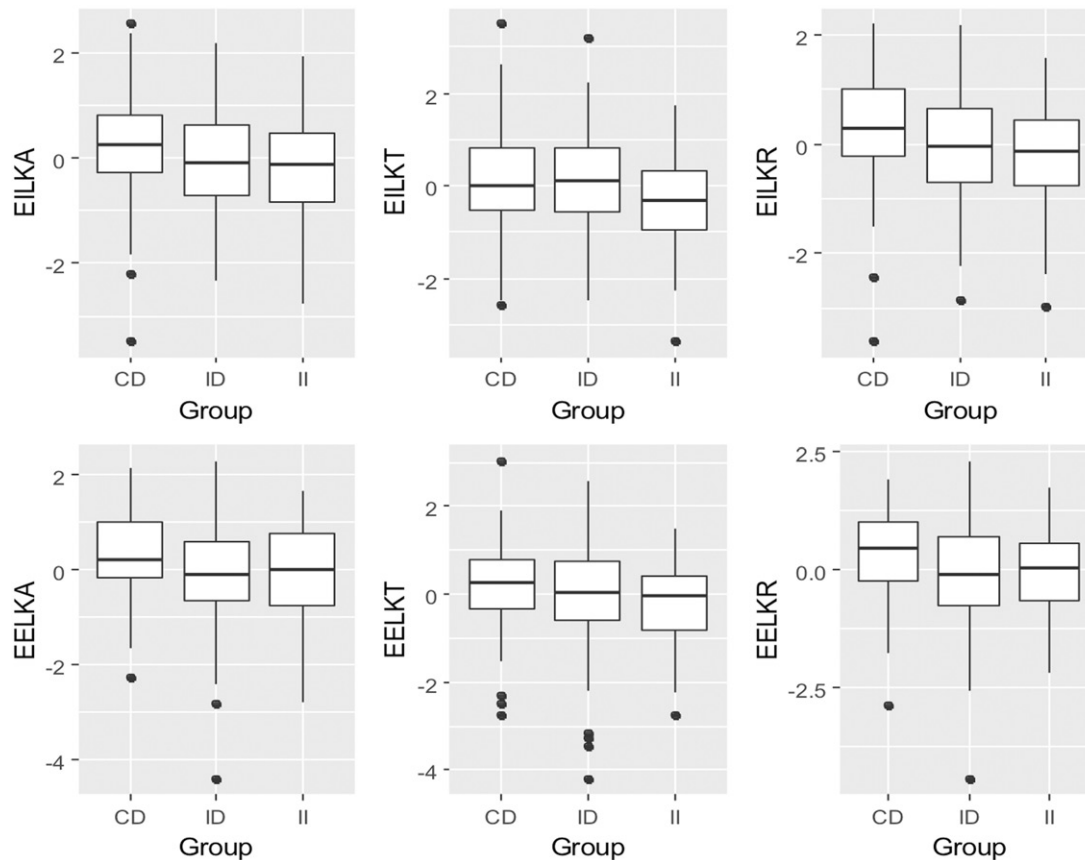


Figure 2. The efficiency of different instructions per group. CD: collaborative deductive; ID: independent deductive; II: independent inductive; EELKA: efficiency (extraneous load and knowledge acquisition); EILKA: efficiency (intrinsic load and knowledge acquisition); EELKR: efficiency (extraneous load and knowledge retention); EILKR: efficiency (intrinsic load and knowledge retention); EELKT: efficiency (extraneous load and knowledge transfer); EILKT: efficiency (intrinsic load and knowledge transfer). The central horizontal line represents the median; the top and bottom of the boxes represent the 25th and 75th percentile. The lines on top and below the boxes extend no further than 1.5 times the interquartile range outside the box. Data points further away are considered as outliers and are indicated as dots.

Table 2. Comparison of instructional methods and effects on different types of cognitive load.

	CD -> ID	CD -> II	Chi-sq, Df	p value
IL	0.116 ± 0.026	0.262 ± 0.036	48.61, 2	<0.001*
EL	0.147 ± 0.026	0.230 ± 0.035	39.89, 2	<0.001*
GL	-0.086 ± 0.021	-0.202 ± 0.029	47.68, 2	<0.001*

*Significant as *p* value is <0.05.

CD: collaborative deductive; ID: independent deductive; II: independent inductive.

Table 3. Comparison of instructional methods and effects on different scores of instructional efficiency.

	CD -> ID	CD -> II	Chi-sq, Df	p value
EELKA	-0.642 ± 0.127	-0.845 ± 0.186	18.13, 2	0.0012*
EILKA	-0.503 ± 0.131	-0.913 ± 0.178	19.01, 2	0.0001*
EELKR	-0.619 ± 0.146	-0.742 ± 0.196	14.10, 2	0.0008*
EILKR	-0.501 ± 0.141	-0.835 ± 0.189	14.29, 2	0.0008*
EELKT	-0.356 ± 0.156	-0.775 ± 0.211	8.59, 2	0.0136*
EILKT	-0.227 ± 0.156	-0.848 ± 0.212	14.08, 2	0.0008*

*Significant as *p* value is <0.05.

CD: collaborative deductive; EELKA: efficiency (extraneous load and knowledge acquisition); EILKA: efficiency (intrinsic load and knowledge acquisition); EELKR: efficiency (extraneous load and knowledge retention); EILKR: efficiency (intrinsic load and knowledge retention); EELKT: efficiency (extraneous load and knowledge transfer); EILKT: efficiency (intrinsic load and knowledge transfer); ID: independent deductive; II: independent inductive.

VPs with lecture, which are collaborative deductive, independent deductive, and independent inductive.

The findings show that the collaborative deductive approach was the most efficient instructional design with significantly higher scores in knowledge acquisition, retention and transfer, as compared with independent deductive and independent inductive approaches. Furthermore, the collaborative deductive group experienced significantly lower intrinsic and extraneous load and higher germane load when compared to the independent deductive and independent inductive groups. Our results agree with previous work that showed an inverse relationship between transfer test performance and intrinsic and extraneous cognitive load for the students who learned in groups (Kirschner et al. 2009b).

It was mentioned that novice learners experienced a high cognitive load when handling complex tasks with several interacting elements (Van Merriënboer and Sweller 2010). However, such high cognitive load can be reduced by sharing it among learners as in the case of collaborative learning, because processing of information is divided across a larger reservoir of their cognitive capacity (Kirschner et al. 2009a). That is why collaborative learning, in our study, resulted in deeper processing of the information, more meaningful learning, and eventually better transfer of knowledge (Kirschner et al. 2009b). Therefore, our hypothesis is completely supported by the attained results.

Collaborative learning can take various forms such as in face-to-face problem-based learning, or synchronous or asynchronous computer-mediated environments (Crook 1996). Our study involved students' collaboration while seated together in the classroom (face to face), but the virtual patient displayed by the computer data acted as a trigger for discussion. Working with computer-based simulated cases in small groups is not enough to increase the scope and depth of the elaboration of computer cases (de Leng et al. 2009). Tutors, therefore, should ensure the harmony

among learners without having dominant or dormant members in the group.

Our study showed a significantly lower efficiency of an independent inductive learning approach on knowledge acquisition, retention and transfer. Such findings can be attributed to the high intrinsic load that was imposed on novice learners by complex VP problems. Previous studies showed that complex problems without any prior knowledge of the topic impose high load on working memory and leads to slow and inefficient learning (Clark et al. 2012). Moreover, the situation could be even worse if minimal guidance is provided to novice students, as was the case in the independent inductive approach (Kirschner et al. 2006). The amount of guidance, therefore, should be adjusted as per the expertise of the learner and the complexity of the given problem.

Our findings might disagree with previous work by Alfieri et al. (2013). The authors in their study suggested that presenting principles and rules (theory) after case comparisons might better promote learning than presenting the principles before case comparisons. In case comparisons, the tutor asks the students to find similarities and differences between cases before encountering the theory. In our context, the students encounter patient problems (VPs) that require them to reach a diagnosis and treatment plan, which is different than finding similarities and differences between two worked examples. Therefore, deductive methods seem to be superior to inductive methods under some conditions but may be inferior under other conditions.

The study showed a negative correlation between intrinsic and extraneous load on one side and germane load on the other side. Such findings are consistent with the principles of cognitive load theory. Imposing high intrinsic and extraneous load results in leaving no or minimal cognitive resources. Working memory resources (germane load) are necessary to process elements related to previous tasks or to link knowledge already available in long-term memory with new knowledge leading to the construction of cognitive schemata (Van Merriënboer and Sweller 2010).

Our study is not without limitations. Due to the limited number of students per batch, we had to extend the study over two years on two different batches of students. Therefore, a complete randomization for all the students together was impossible to achieve. This probably also was a reason for not optimal fit in the models for the efficiency measures. The study has also hypothesized the inductive learning approach to presenting the VP (problem) before the lecture (direct instruction). In general, sequencing VP in an inductive or deductive approach should carry differences in the amount of guidance and complexity of both the VP and the lecture to optimally scaffold learning. However, our aim was to ensure consistency in the presented content between the inductive and deductive groups. Furthermore, the collaborative deductive method involved a tutor, which might be counted as a variable that is not present in the other two groups. However, we cannot isolate the tutor from the collaborative activity as it represents one of the elements of the collaborative activity in our context. The tutor role was more as a facilitator of the collaborative session than being a knowledge provider.

The study has provided some practical implications for both medical teachers and VPs designers. For medical teachers, the study has shown that the most efficient

instructional method is a collaborative deductive approach: first present relevant theory in a lecture and then let students work collaboratively with the virtual patients. The study has confirmed previous work grounded on cognitive load theory that declared the significance of adapting the amount of guidance based on the complexity of the provided problem and the experience of the learners. Furthermore, the study has shown that any future selection of instructional methods involving VPs should consider a cognitive load measurement, as some VPs applications might be detrimental to learning.

The study has also provided a practical implication in regard to VPs design specificity: the VP design should be customized during development based on its intended use by the teacher and students. VP designers should ask medical teachers “how they are going to use the VPs?” before commencing the development process. A collaborative use of VPs requires a more complex design to trigger discussion and reflection during use, while an inductive independent use of VPs requires relatively simple VPs that are able to stimulate rather than hinder learners to close knowledge gaps that might have arisen during VP practice.

Conclusions

Collaborative use of VPs after a lecture is the most efficient instructional method, of those tested, as it leads to better learning and transfer combined with lower cognitive load, when compared with independent use of VPs, either before or after the lecture.

Ethical approval

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

Notes on contributors

Hesham F. Marei, MSc, FDSRCS (Eng.), PhD (OMFS), MHPE, is an Associate Professor and Consultant of Oral and Maxillofacial Surgery, at College of Dentistry, Imam Abdulrahman Bin Faisal University, Saudi Arabia. Dr. Marei is currently a PhD fellow at School of Health Professions Education, Maastricht University, the 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, the Netherlands.

Mohamed M. Al-Eraky, MBBCh, MSc, MMed, PhD, is an Assistant Professor of Medical Education. He is currently appointed as Director of Development & Academic Initiatives at the Vice-President Office for Academic Affairs at Imam Abdulrahman Bin Faisal University, Saudi Arabia.

Jeroen J. G. Van Merriënboer, 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, the Netherlands. He is Research Director of the School of Health Professions Education.

Acknowledgements


Authors would like to thank all the members of Oral and Maxillofacial Surgery Division who revised the lectures content and established the content validity of the different tests.


Disclosure statement


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

ORCID

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

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

Mohamed M. Al-Eraky  <http://orcid.org/0000-0003-2015-7630>

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

References

- Alfieri L, Nokes-Malach T, Schunn C. 2013. Learning through case comparisons: a meta-analytic review. *Educ Psychol.* 48:87–113.
- Bateman J, Allen M, Samani D, Kidd J, Davies D. 2013. Virtual patient design: exploring what works and why. A grounded theory study. *Med Educ.* 47:595–606.
- Bates D, Mächler M, Bolker B, Walker S. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823.*
- Berman N, Fall LH, Smith S, Levine DA, Maloney CG, Potts M, Siegel B, Foster-Johnson L. 2009. Integration strategies for using virtual patients in clinical clerkships. *Acad Med.* 84:942–949.
- Brunken R, Plass JL, Leutner D. 2003. Direct measurement of cognitive load in multimedia learning. *Educ Psychol.* 38:53–61.
- Clark R, Kirschner PA, Sweller J. 2012. Putting students on the path to learning: The case for fully guided instruction. *American Educ.* 36:6–11.
- Crook C. 1996. Computers and the collaborative experience of learning. Psychology Press.
- de Leng BA, Muijtjens 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.
- Edelbring S, Brostrom O, Henriksson P, Vassiliou D, Spaak J, Dahlgren LO, Fors U, Zary N. 2012. Integrating virtual patients into courses: Follow-up seminars and perceived benefit. *Med Educ.* 46:417–425.
- Ellaway R, Candler C, Greene P, Smothers V. 2006. An architectural model for MedBiquitous virtual patients. Baltimore, MD: MedBiquitous.
- Ellaway R, Davies D. 2011. Design for learning: Deconstructing virtual patient activities. *Med Teach.* 33:303–310.
- Ellaway R, Topps D, Lee S, Armson H. 2015. Virtual patient activity patterns for clinical learning. *Clin Teach.* 12:267–271.
- 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.
- Huwendiek S, Reichert F, Bosse HM, de Leng BA, van der Vleuten CP, Haag M, Hoffmann GF, Tonshoff B. 2009. Design principles for virtual patients: a focus group study among students. *Med Educ.* 43:580–588.
- Kirschner F, Paas F, Kirschner PA. 2009a. A cognitive load approach to collaborative learning: United brains for complex tasks. *Educ Psychol Rev.* 21:31–42.
- Kirschner F, Paas F, Kirschner PA. 2009b. Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Comput Human Behav.* 25:306–314.
- Kirschner PA, Sweller J, Clark RE. 2006. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ Psychol.* 41:75–86.
- Leppink J. 2016. Cognitive load measures mainly have meaning when they are combined with learning outcome measures. *Med Educ.* 50:979.

- Leppink J, Paas F, Van der Vleuten CP, Van Gog T, Van Merriënboer JJ. 2013. Development of an instrument for measuring different types of cognitive load. *Behav Res Methods*. 45: 1058–1072.
- Leppink J, Paas F, van Gog T, van der Vleuten CPM, van Merriënboer JJG. 2014. Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learn Instr*. 30:32–42.
- Marei HF, Donkers J, Al-Eraky MM, van Merriënboer JJG. 2017. The effectiveness of sequencing virtual patients with lectures in a deductive or inductive learning approach. *Med Teach*. 39:1268–1274.
- Paas FG, Van Merriënboer JJ. 1993. The efficiency of instructional conditions: an approach to combine mental effort and performance measures. *Hum Factors*. 35:737–743.
- Team RC. 2017. A language and environment for statistical computing. R foundation for statistical computing, 2012.
- Tremblay ML, Lafleur A, Leppink J, Dolmans DH. 2017. The simulated clinical environment: cognitive and emotional impact among undergraduates. *Med Teach*. 39:181–187.
- Van Gog T, Paas F. 2008. Instructional efficiency: revisiting the original construct in educational research. *Educ Psychol*. 43:16–26.
- Van Merriënboer JJ, Sweller J. 2010. Cognitive load theory in health professional education: Design principles and strategies. *Med Educ*. 44:85–93.
- Van Merriënboer JJG, Kirschner PA. 2018 *Ten steps to complex learning* (3rd Rev. Ed.). New York, Routledge.
- Vandewaetere M, Manhaeve D, Aertgeerts B, Clarebout G, Van Merriënboer JJ, Roex A. 2015. 4c/id in medical education: How to design an educational program based on whole-task learning: Amee guide no. 93. *Med Teach*. 37:4–20.
- Young JQ, Van Merriënboer J, Durning S, Ten Cate O. 2014. Cognitive load theory: implications for medical education: Amee guide no. 86. *Med Teach*. 36:371–384.