The Effectiveness of Educational Applications: Two Small Empirical Investigations

Italo Masiello*,1, Kirsti Lonka[1,2], Matti Nikkola[3] and Helga Hirschfeld[4]

[1]*Department of Learning, Informatics, Management and Ethics, Centre for Medical Education, Karolinska Institutet, Sweden*

[2]*Department of Applied Science of Education, University of Helsinki, Finland*

[3]*Department of Cellular and Molecular Biology, Karolinska Institutet, Sweden*

[4]*Department of Neurobiology, Care Sciences and Society, Division of Physiotherapy, Karolinska Institutet, Sweden*

Abstract: ICT is thought to be successful in helping students to learn more deeply and actively. Despite its popularity, evidence as to its effectiveness is undergoing increased scrutiny. This article reports on the empirical investigation of the implementation of two interactive animations. Questionnaires measured students' learning styles, ICT competence and students' impressions of the animations. In Study I, medical students were matched-paired into experimental and control groups, and their performance was compared by a short written test. In Study II, physiotherapy students were arranged into two pre-existing groups. The results in Study I showed that there was no statistical difference between the control and experimental groups. A positive trend was found between *deep processing* and general test score. Also statistical significance was found between *deep processing* and three-dimensional perception. In Study II, the two groups performed equally on a final exam, but the experimental group scored significantly higher on a specially constructed question to test the effect of the animation. Further, *Lack of Regulation* had a negative effect on the scores of a final exam. The fundamental result of these small studies was that of deep learning and a lack of constructive regulative study patterns being more relevant for the learning tasks than the interactive applications themselves. This has implications for teachers who should explore alternative teaching strategies to promote deep learning.

Keywords: Interactive learning environments, educational applications, deep approach.

INTRODUCTION

Many educators share the covert assumption that information communication technology (ICT) can help students of any domain to learn more effectively [1]. Medicine is not an exception, and because medical education is a complex and ill-defined domain, the main difficulty is to find an efficacious way of conveying the large amount of background knowledge necessary for the practice of medicine [2]. A correct implementation of ICT should, at least in theory, stimulate deep and active learning and understanding, rather than superficial and passive information processing. The deep approach to learning is characterized by an intention to understand material, relating ideas to one's previous knowledge, and relating evidence to conclusions [3]. The surface approach is characterized by an intention simply to reproduce parts of a content, passive acceptance of information and ideas and memorizing facts routinely [3]. Educational research affirms that the approaches the students embrace can predict how they will learn and succeed in their studies [3-6]. Students who adopt a deep approach are able to regulate their own learning [7] and more likely to adapt to different learning environments. However, there is also recent and contrasting evidence that in general no powerful predictors of learning in higher education have been identified [8].

The ability to actively interact with learning applications is claimed to positively influence user attitudes towards the learning environment by enhancing learners' motivation [9]. All this should be due to multi-sensorial engagement. Also, according to cognitive-load theory, information presentation techniques should be designed to ease the load on working memory and consequently facilitate learning through mental model acquisition and automation [10]. Indeed, 3D Embryo and Movement in Space were produced with the idea of facilitating the creation of mental models of two different and complex phenomena. However, based on the fact that technology alone is not a sufficient condition to enhance learning [11], we wanted to test the extent of the effect of deep approach and the interplay between previous IT skills and the interactive multimedia.

THE TWO APPLICATIONS

3D Embryo

3D Embryo is an interactive animation of a specific period of the human embryonic development. It was produced by two medical students, Hanna Reuterborg and David Örtoft, to facilitate the understanding of early embryology. Their idea behind the multimedia application was to simplify complex medical processes in an innovative way. 3D

*Address correspondence to this author at the Berzelius väg 3, 17177 Stockholm, Sweden; Tel: +46 (8) 524 87856; Fax: +46 (8) 345128; E-mail: italo.masiello@ki.se

Embryo graphically elucidates what happens at overlapping times when several layers of tissue develops into differentiated organs and parts of the more familiar foetus. The application provides short explanations of concepts and items related to the animation but does not provide though any performance support in terms of self-test.

Movement in Space

Movement in Space is an interactive animation that simulates a person's movements within a reference structure in space—the Cartesian coordinate system (http://mediabyran.kib.ki.se/animations/movement.html). It was produced and developed by Thomas Nixon and Helga Hirschfeld who were concerned about the difficulty that physiotherapy students have in understanding machine-generated graphs for a correct diagnosis of the patient's disability. *Movement in Space* graphically explicates how human movement is accomplished by the use of terms position, velocity and acceleration, kinematics and biomechanics definitions. The application provides a self-test module to evaluate the understanding of these aforementioned definitions.

METHODS

Study I

Participants and Procedure

The participants were first-year medical students taking an undergraduate cell biology course, which included embryology. Ninety-two out of about 130 students attending the course elected to participate in the study and were informed of the rights as participants, according to Swedish Medical Research Council guidelines. Only complete data from 88 (67%) participants could be used. Fifty-seven were females and 31 males.

The research study was randomized matched-pairs experimental design. At the outset, the participants were divided into two clusters that are as similar as possible on the characteristics of previous knowledge of the subject of embryology, entrance basis to the medical program and gender, information that originated from the department and a questionnaire that students answered beforehand. Then, we randomly choose one of the two clusters to be the experimental group and the other to be the control.

The two groups of students followed the same periods throughout the course up until the developmental biology period, and more specifically the early embryology part of that period. Subsequently, for one day the groups followed separate tracks. The experimental group was taught by aid of slides and 3D Embryo, whereas the control group without the use of 3D Embryo. The students in the experimental group had a further 2-hour study session in a computer laboratory using 3D Embryo. After a week from the lectures and lab both groups were tested on the subject of early embryology. The control group did not experience 3D Embryo only until after the test.

MATERIAL AND DATA

The Inventory of Learning Styles and Students' Impressions Questionnaires

A few weeks before the beginning of the one-day separate tracks program the students answered a questionnaire that investigated the students' learning styles and attitudes to ICT [12]. The instrument measured *deep processing*, *stepwise processing*, Self-regulation, and Lack of Regulation [13]; Time Management [14]; and IT Competence [12]. The *deep* and *stepwise processing* scales are similarly to the *deep and surface approaches* to studying dimension of Marton and Säljö [3].

The reason for adopting such a combined approach is related to the way in which aspects of the learning styles could easily be linked to various characteristics of the learning environment and the application, 3D Embryo. The responses were structured as a five-point Likert scale ranging from 5 (Highly Agree) to 1 (Highly Disagree), keeping with the original format.

The test

A week after the lecture and the lab, the students took a short paper-and-pencil test to assess the effects of 3D Embryo as a learning tool. The test assignment was composed of questions designed to examine various aspects of understanding of the subject of early embryology: *what*, the ability to retrieve factual information, (2 questions); *why*, the ability to contextualize and conceptualize information, (2 questions); *3D/visualization,* use of spatial abilities, (2 questions); and *parallel process*, the ability to relate several events starting and ending at different times but occurring concurrently, (2 questions); a total of eight questions.

Statistical Analysis

We used basic statistics, box plots, to test the homogeneity of the two groups on several factors, and a t-test compared the test scores between the two groups. One-way ANOVAs tested the significance of each of the scale of the inventory. Spearman rank order correlation found out the relations between factors. And two-way ANOVAs were used to test main effects between factors.

Study II

Participants and Procedure

The participants were second-year physiotherapy students taking an undergraduate course in Kinesiology, which included kinematics theory and movement analysis with respect to a three-dimensional reference system in space. All students (61) in the course participated in the study, and were informed of their rights as volunteer participants. However, we collected complete data of only 44 (72%) students. Twenty-six were females and 18 males.

The research study we utilized was a cohort design. Being such, we could not control for any variance in or homogeneity of the groups. The participants were two pre-existing study groups and each was randomly assigned to be either control or experimental group. The two groups of students followed the same periods throughout the course up until the laboratory session. There, they both would practice hands-on movement analysis with traditional instruments but in two distinct periods in order to make available the interactive simulation only to the experimental group. Both groups received the same amount of instruction and lecture time. Ten days after, both groups took a written examination.
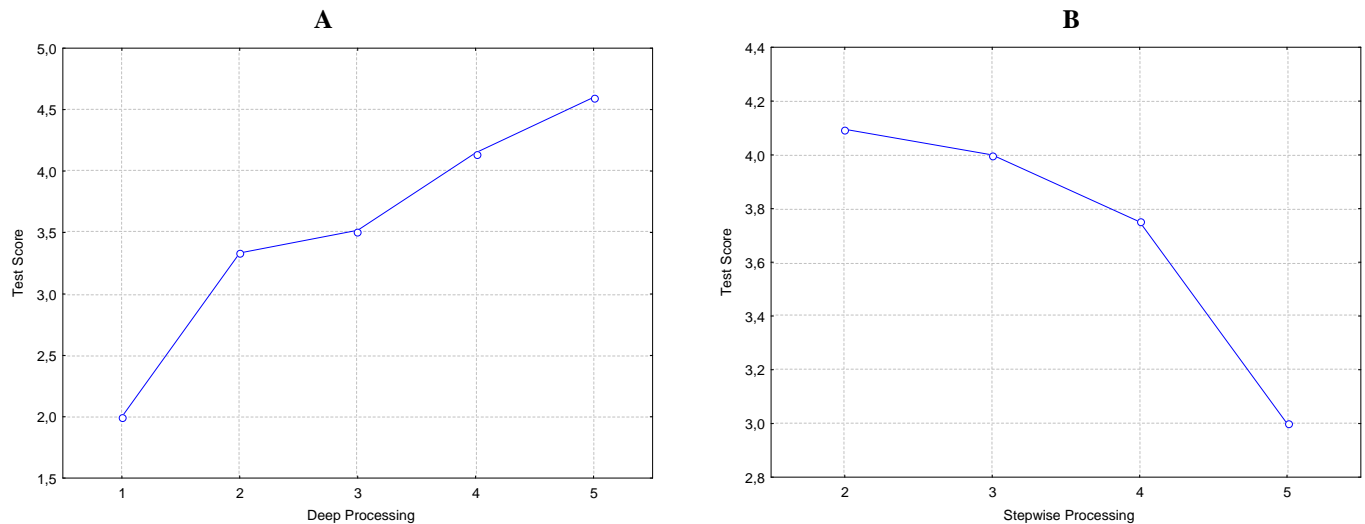
A

B

Fig. (1). The two graphs show the results of the one-way ANOVAs of the *deep* (**A**) and *stepwise* (**B**) *processing* scales when compared against the test scores. The average test score is presented on the Y-axis, while the processing scales are shown on the X-axis.

MATERIAL AND DATA

The Inventory of Learning Styles

As in study I, in the present study we combined several instruments to measure different aspects of our research question. A few days before the laboratory session the students answered a questionnaire that investigated the students' learning styles and attitudes to ICT. The instrument used was the same as in Study I (see Section *Material and Data* in Study I), as well as the format and the reason for adopting it.

The Examination

The examination the students took was in pencil-and-paper format and contained a blend of short-essay, multiple-choice and short-phrase answers about general physiotherapy knowledge. One question, divided in four parts and giving 0-4 points, tested the understanding of movement analysis with respect to the three dimensional coordinate system in space and compared the two groups on the outcome measure.

Statistical Analysis

We carried out the same analyses as in Study I. A t-test for independent groups compared the exam scores and the movement analysis question between the two groups. Spearman rank order correlation found out the relations between factors. One-way ANOVAs were used to test significance between the scales, two-way ANOVAs tested main effects, and a factorial ANOVA tested the interaction.

RESULTS

Study I

The comparison of the test scores of the two groups showed that on average the experimental group scored slightly higher than the counterpart, but the difference was not statistically significant.

One-way ANOVAs of the *deep processing* scale showed a positive trend. As the use of the *deep processing* learning style increased (from 1 to 5), the tendency of a higher score

in the test also increased ($F_{(4,83)} = 2.383$, $p = 0.057$). This is also inversely proportional to the *stepwise* scale, which decreases with higher test scores ($F_{(3, 84)} = 0.551$, $p = 0.648$) (Fig. **1**).

Two-way ANOVA showed main effects of the *deep processing* scale against the visualization question ($F_{(4,82)} = 2.668$, $p = 0.037$). The students who scored higher on the *deep processing* scale also obtain at least one right answer on average on the visualization questions (Fig. **2**).
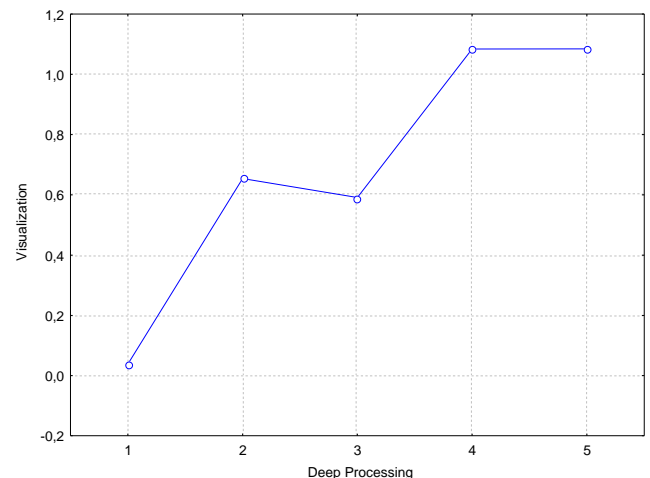

Fig. (2). The graph represents the result of a two-way ANOVA. On the X-axis it is represented the *deep processing* learning style, while effect on the average score on the visualization question is showed on the Y-axis.

We also observed a main effect for entrance to medical program ($F_{(2,82)} = 3.456$, $p = 0.036$); on average those entering the program by means of an exam scored higher on our test than those entering *via* interviews or grades (Fig. **3**).

In addition we carried out Spearman correlation to check of any relations between the factors considered (Table **1**). A strong, positive correlation was found between Self-
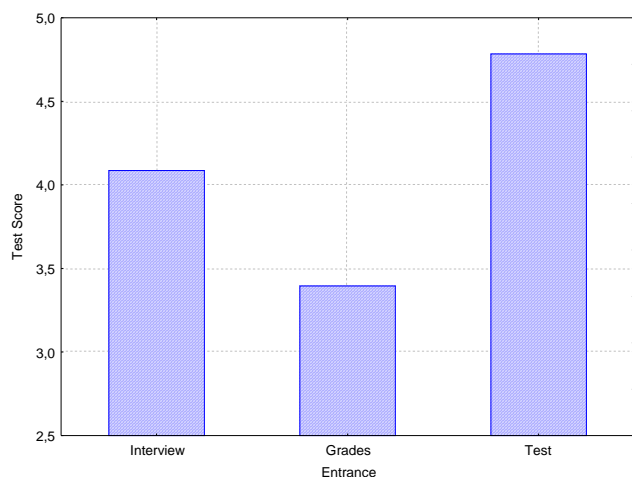
Fig. (3). This graph represents the main effect on test scores for means of entrance to the medical program. The graph represents the average test score, on the Y-axis, distributed along the various entrance bases to the program-interview, grades and test-showed on the X-axis.

regulation and *deep processing* scales (R = 0.63). Other week and positive correlations were found between *stepwise processing* and Time Management (R = 0.35), *deep processing* and test scores (R = 0.34), and *deep processing* and visualization question (R = 0.31).

Study II

When comparing the exam scores of the two groups with a t-test, the experimental group scored higher (63.1) than the control group (58.9) on average. Although we saw a positive trend, the difference was not statistically significant (t(42) = -0.71, p = 0.478). The box plot and t-test of the two groups

compared on the movement analysis question showed instead the experimental group scored statistically higher (t(42) = -2.043, p = 0.047).

ANOVA analysis showed that higher marks on the scale Lack of Regulation corresponded to lower exam scores (F(2,40) = 3.678, p = 0.034). A factorial ANOVA was used to test the main effects and interaction of IT Competence and the score on the movement question. In this case, the control group who on average scored lower marks on the IT Competence scale scored higher on the movement question (F(2,38) = 4.155, p = 0.023) (Fig. **4**). There was also interaction effect; on average the experimental group who revealed ICT skills scored higher than the control group on the movement analysis question (F(2,38) = 8.700, p = 0.0007).

Spearman correlations analysis revealed a positive correlation between Exam score and Movement analysis question (R = 0.56) (Table **2**). We also found weak negative correlations between the scales l*ack of regulation* and *deep processing* against the exam score factor (R = -0.32, respectively for both), between the groups and the factor traditional learning (R = -0.33), and between IT Competence and the factor traditional learning (R = -0.38). IT Competence and Movement analysis factors (R = -0.40) as well showed a negative correlation. Also a positive and week correlation was found between groups and the Movement question (R = 0.30).

DISCUSSION

Study I

Our results show that the interactivity and third dimension of 3D Embryo did not help the students in getting higher marks on our test when compared to students who did not use the three-dimensional tool. As Garg *et al*. [15] stated "…these advantages may be more imagined than real."

Table 1.   Spearman Rank Order Correlations, Study I

	Groups	**Entrance**	**Visuali-sation**	**Stepwise**	**Lack of Reg.**	**Time Manag**	**Deep**	**Self-Reg.**	**IT Comp.**	**Trad. Learn.**	**Test Scores**
Groups	1.00										
Entrance	-0.17	1.00									
Visualization	-0.07	0.06	1.00								
Stepwise	0.03	0.07	-0.11	1.00							
Lack of Reg.	0.06	-0.07	-0.21	0.15	1.00						
Time Manag.	0.09	0.02	0.02	*0.35	*-0.36	1.00					
Deep	-0.10	0.01	*0.31	0.07	-0.12	0.24	1.00				
Self-Reg.	0.04	-0.11	0.22	0.19	-0.05	0.29	*0.63	1.00			
IT Comp.	-0.13	-0.19	-0.04	0.03	-0.05	0.04	-0.00	0.13	1.00		
Trad. Learn.	-0.09	0.17	-0.06	0.09	-0.07	-0.02	-0.11	-0.02	-0.25	1.00	
Test Scores	-0.06	-0.03	*0.59	-0.11	-0.20	0.05	*0.34	0.10	0.07	-0.03	1.00

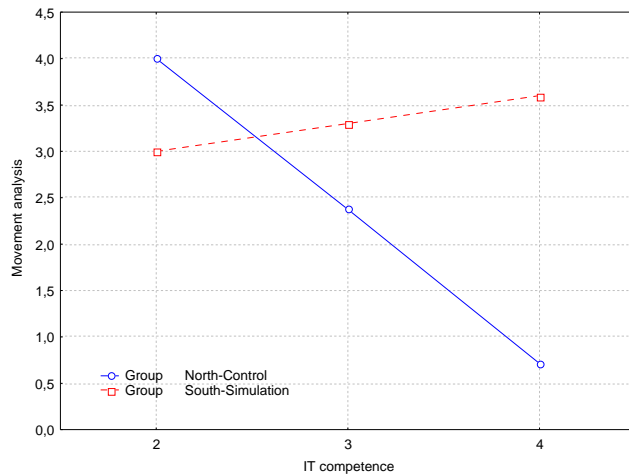*Marked correlations are significant at p < 0.05.

Fig. (4). The graph represents the factorial ANOVA used to test the main effects of IT Competence (X-axis) and groups (dotted line = control group; full line = experimental group) on the score of the Movement analysis question (Y-axis) and interaction between the groups and IT Competence.

Instead, the *deep processing* scale of the learning styles questionnaire showed an interesting trend. The results indicated that the students who adopted a *deep processing* learning style were able to score higher marks than those who adopted *stepwise processing*, no matter the use of a three-dimensional aiding application. Moreover, the students adopting the *deep processing* style also obtained on average one out of the two visualization questions right, again regardless the use of 3D Embryo.

An interesting outcome was that of higher test scores on average for those students entering the medical program through a general exam. This result is in contrast to that of a local report [16] for which those students admitted on the basis of an aptitude test are lower achievers. However, this result must be taken with caution since a small test like the one used in this study cannot possibly explain the complexity of recruitment to the medical programme.

The research instrument confirmed the conventional coupling of the Self-regulation and *deep processing* scales demonstrated by the strong correlation coefficient between these factors, showing consistency and reliability of the scales. Furthermore, previous ICT competence did not have any influence on learning from a simulation.

Study II

In Study II, contrary to the previous, we had the opportunity to analyse the data from the final exam of the course. Clearly, the final exam did not only test movement analysis but also other aspects of physiotherapy taught in the course. However, one question, divided in four possible parts, tested the analysis of human movement accomplished by the use of terms position, velocity and acceleration. In regards to the results on the overall test, the two groups of students performed similarly, although they differed on the score of the movement analysis question, with the experimental group scoring statistically higher than the control group.

In this study, neither *deep* nor *stepwise processing* seems to have played a part in the students' assimilation of the subject. Lack of Regulation, often associated with surface approach to learning, proved instead to have mediated test scores. Indeed, students who revealed a lack of regulative mechanisms in their studying also scored much lower in the overall test. This is to reinforce the concept that *stepwise processing*, surface approach, is a dysfunctional pattern of studying.

The analysis of ICT competence and the groups on the movement analysis question showed some effects. Previous ICT skills did not help the scores on the movement analysis question. Actually, students of both groups who revealed a lack of ICT skills scored higher marks on this question and were able to get on average between 3-4 points, out of possible 4, which may be somewhat counterintuitive.

Table 2.   Spearman Rank Order Correlations, Study II

	Groups	**Exam Score**	**Movement Analysis**	**Deep**	**Surface**	**Strategic**	**Time Manag.**	**Lack of Reg.**	**Trad. Learn.**	**IT Comp.**
Groups	1.00									
Exam score	0.11	1.00								
Movement Analysis	*0.30	*0.56	1.00							
Deep	-0.07	-0.15	-0.22	1.00						
Surface	0.10	-0.03	-0.21	0.01	1.00					
Strategic	0.00	0.07	0.07	0.28	-0.14	1.00				
Time Manag.	0.17	-0.24	-0.23	-0.16	0.22	0.12	1.00			
Lack of Reg.	-0.06	*-0.32	-0.19	*-0.32	-0.02	-0.25	0.12	1.00		
Trad. Learn.	*-0.33	-0.05	-0.01	-0.20	0.12	-0.21	-0.12	0.03	1.00	
IT Comp.	-0.06	0.02	*-0.40	0.07	0.22	0.11	0.28	-0.20	*-0.38	1.00

*Marked correlations are significant at $p < 0.05$

GENERAL DISCUSSION

There is a limitation to the studies, that is, the short time intervention. The use of the simulation in the lab only lasted a few hours. It is hard to acknowledge that a few-hour session may contribute to a lasting experience. As Rouet [17] and Goldman [18] suggest, adequate training is necessary to benefit from complex displays and computer technologies. There are other limitations as well, like students' motivation and constraints of real-life educational settings and the possible influence played by the two extra hours of study that the experimental group in Study I received. However, given all the limitations, the majority of students found the applications interesting and fun to use, and stated that they would like to use more such learning tools.

CONCLUSIONS

To date the advantage of educational technology over traditional education is not always self-evident [19, 20]. Our two studies confirmed this notion. The studies showed that students using a more traditional method of learning and teaching scored similarly in a performance test when compared to students who used a "modern" method aided by an interactive educational tool.

Nonetheless, the results from these studies demonstrated and confirmed that deep approach is preferred over dysfunctional studying patterns in order for learners to succeed in their study, no matter the use of technology. This research cannot provide generalizable results, but it can suggest that teachers should pay attention to the way the students approach their learning since that plays a significant part in the learning process.

REFERENCES

[1]   Harden RM. Evolution or revolution and the future of medical education: Replacing the oak tree. Medical Teacher 2000; 22(5): 435-42.

[2]   Custers EJFM, Boshuizen HPA. The psychology of learning. In: Norman G, van der Vleuten CPM, Newble DI, Eds. International Handbook of Research in Medical Education. Dordrecht, The Netherlands: Kluwer Academic Publishers 2002. p. 163-204.

[3]   Marton F, Säljö R. On qualitative differences in learning. I - Outcome and Process. Br J Educ Psychol 1976; 46: 4-11.

[4]   Lindblom-Ylänne S, Lonka K. Individual ways of interacting with the learning environment -- are they related to study success? Learn Instruct 1999; 9(1): 1-18.

[5]   Entwistle N, Ramsden P. Understanding Student Learning. London, UK: Croom Helm 1983.

[6]   Vermunt JDHM, van Rijswijk FAWM. Analysis and development of students' skill in selfregulated learning. Higher Educ 1988; 17: 647-82.

[7]   Vermunt JD, Vermetten YJ. Patterns in student learning: Relationships between learning strategies, conceptions of learning, and learning orientations. Educ Psychol Rev 2004; 16(4): 359-84.

[8]   Coffield F, Moseley D, Hall E, Ecclestone K. Learning styles and pedagogy in post-16 learning. A systematic and critical review. 2004 [cited 2007 January]; Available from: http: //www.lsda.org.uk/files/PDF/1543.pdf

[9]   Kettanurak VN, Ramamurthy K, Haseman W. User attitude as a mediator of learning performance improvement in an interactive multimedia environment: an empirical investigation of the degree of interactivity and learning styles. Int J Hum-Comput Stud 2001; 54(4): 541-83.

[10]   Tindall-Ford S, Chandler P, Sweller J. When two sensory modes are better than one. J Exp Psychol 1997; 3(4): 257-87.

[11]   Masiello I. Learning in Focus - Rethinking the Role of Technology in Medical Education [Doctoral Thesis]. Stockholm: Karolinska Institutet 2005.

[12]   Masiello I, Ramberg R, Lonka K. Attitudes to the application of a Web-based learning system in a microbiology course. Comput Educ 2005; 45(2): 171-85.

[13]   Vermunt JD. Inventory of learning styles in higher education. The Netherlands: Leiden University 1994.

[14]   Tait H, Entwistle NJ, McCune V. ASSIST: A reconceptualization of the Approaches to Studying Inventory. In: Rust C, Ed. Improving Student Learning: Improving Students as Learners. Oxford, UK: Oxford Brookes University 1998; p. 262-70.

[15]   Garg AX, Norman GR, Eva KW, Spero L, Sharan S. Is there any real virtue of Virtual Reality? The minor role of multiple orientations in learning anatomy from computers. Acad Med 2002; 77(10S): S97-9.

[16]   Lonka K, Hindbeck H, Scheinin P. Uppföljning av lokal antagning till läkarutbildning vid Karolinska Institutet - genomströmning och studiepresentation för studerande antagna vt1997-vt2001. Forskningrapport. Stockholm: Karolinska Institutet 2004 May.

[17]   Rouet J-F. Designing multimedia systems for learning: Some lessons and further issues. In: Rouet J-F, Levonen J, Biardeau A, Eds. Multimedia Learning Cognitive and Instructional Issues. Oxford, UK: Elsevier Science Ltd 2001; pp. 167-72.

[18]   Goldman SR. Issues for deeper understanding of visual learning processes and representation. Eur J Psychol Educ 1999; 14: 295-300.

[19]   Goldman SR. Learning in complex domains: When and why do multiple representations help? Learn Instruct 2003; 13: 239-44.

[20]   Makridou-Bousiou D. The effectiveness of technology in teaching high school Economics. J Information Technol Impact 2006; 6(1): 9-18.