Evaluation of 3D environments and virtual realities in science teaching and learning: The need to go beyond perception referents

Kunle Oke Oloruntegbe and Gazi Mahabubul Alam*

Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia.

Accepted 1 April, 2010

There is a raging debate in literature about the effectiveness of 3D environments and virtual realities. The debate seems an unending one with both the advocates and critiques striving to justify their various positions. A survey of literature, however, revealed the absence of sufficient data to justify taking a position. This led the authors to hypothesize that most of the evaluation carried out in this area did go beyond perception referents. Over one thousand articles in reputable journals and proceedings were consulted including a few others without affiliation. Twenty four that covered the amorphous subjects of evaluation, assessment, effectiveness, potentials and benefits were painstakingly analyzed for their coverage particularly over the three learning outcomes of cognitive, skills and attitudes. The findings revealed that while there were evidences of learning improvement and performance enhancement, over eighty five percent of the data revolved around affective dispositions of satisfaction, interest, enjoyment and fun. With this, the authors concluded that there was no sufficient evidence yet to back up the pedagogical effectiveness of this state-of-art technology. The authors, however, appreciates the newness of this field of human endeavour, and is of the opinion that sufficient time, spread and focus are required to make valid evaluation possible.

Key words: 3D environments, virtual realities, pedagogical effectiveness, learning outcomes.

INTRODUCTION

That 3D visualizations and virtual laboratories have come to stay and have revolutionized science teaching and learning, particularly chemistry, is stating the obvious. What for the numerous advantages that researchers had amplified. There is need to mention a few of them here. They are reported to have improved the attitudes of students toward chemistry and advanced interactive and collaborative teaching and learning processes (Alhalabi et al., 2001; Nejad, 1997). They are useful when physical experiments are too expensive, risky and complex (Matthews, 2002; Dillon, 2007). They enable students to approach chemistry more like practicing scientists and see interesting real-world application of key concepts (Yaron David). They help them carry out pre-work ahead of practical classes, so maximizing the quality of time they spend in the laboratory and help develop excellence in experimental skills (RCS, 2009). They help to banish teachers’ fear about letting students loose in the lab (RSC, 2009); and are effective in modeling abstract concepts and make what is abstract a tangible manipulable concrete (Dalgarno et al., 2003; Strangman et al., 2003). Geographical restriction is no longer a serious problem (Yoshida, 2000). There are some others.

However, a few researchers, that were not so carried away with the fancies of the present day technology, and who would not quickly forget the old school lab practical experiences, have queried if 3D visualizations and interactivity can serve science education better than real-lab hands-on experiences. They have also amplified the shortcomings of 3D environments and highlighted the advantages of brick and mortar, real-laboratories. To them,
many of 3D nature may not be fully grasped on the commonly available flat computer screen thereby generating more misconceptions and ambiguity than reducing them (Canizares and Faut, 1997; Sinex and Gage, 2004; Gervasio, 2004; Tasker and Dalton, 2006). They are not suitable for quick comparison especially when teaching structures in organic chemistry (Gervasio, 2004; Schmidt-Ehrenberg et al., 2002). Beside, in these pedagogical issues, there is also the economic aspect, making the present state-of-art-technology use in science education rather too expensive to bear for many schools, particularly in developing nations (Bon, 2007). The problems of access, poor internet connectivity and digital divide are there, making 3D viewing on computer impossible in many nations’ schools. It was observed that much of the innovative work has been in institutions where systems and tools development has been seeded by project funding (JISC, 2007). Even then, use is often limited to a few departments not to talk of other institutions.

The above arguments point to two main important queries: the educational benefit from learners’ perspectives and the overhead financial implications on the part of the school administrators making the whole exercise a kind of suspect. The question asked by Valdez (2005) “What prevents educational researchers from giving us definitive answers about technology in the classroom that would satisfy both critics and advocates?” is a pertinent one. Also crucial are the verdicts of no conclusive answer to the question whether technology promotes students learning or not (Ma and Nickerson, 2006; Taghavi and Colen, 2009) and the ambivalent feeling and dichotomy among instructors and teachers are reported by Barak (2007). Yet, the ability of these technologies in bridging the gap between concrete world of nature and the abstract world of concepts and models cannot be overlooked. They had been described as more of enablers (Raineri, 2001; Striegel, 2001; Barak, 2007) of learning than of inhibitors (Dewhurst, et al., 2000; Dibiase, 2000). They should be embraced. Riordan (2009) affirms that new technologies replace old ones and no one goes back to the old ones. Word processors have for many replaced pen and paper which in turn replaced quill and parchment. A person could still produce a paper with either of the previous two, the number of people who would go back to quill and parchment will be infinitesimally small. However, before being able to replace the potentials of the new technologies would not be in doubt.

We definitely know that whatever the potency of a new innovation, there are always the other sides of such. It becomes a useful innovation when its advantages far overwhelm the disadvantages. Everyone will stand for such. But the debate about which one is better, at what stage, context and purpose is still an on-going, albeit an unending one. It was like the debate on learning theories that took the central stage decades back, the connectionist tradition, insightful learning, the gestalt and the cognitive theories and so on. There were strong alignments too on resonating expository and inquiry continuum in lesson delivery. One of the authors of this work happened to contribute to such debate, and found out like others that there are conflicting claims to their effectiveness. Such effectiveness was best situated on the stage, on the context and purpose of the methods. While some studies submit that, the inquiry ends foster skill development, retention and transfer of learning are seldomly used by teachers for constrains of time, space, efforts and fund. Others, who are in love with reception learning and expository delivery talk about their effectiveness in the delivery of abstract concepts, and they are widely used by teachers even till today.

The authors would like to look at this present debate in similar context. One would not have taken the submissions on learning theories seriously if they were not arrived at through breathtaking rigorous experimentally designed researches with control. Profound and laudable as the efforts at revolutionizing science learning through 3D world realities are, the question is has the ongoing debate about the effectiveness or superiority of one over the other, between 3D visualizations/virtual realities and real-lab hands-on experiences followed the same pattern of research design and control? It is this question and some others asked earlier that provoked thoughts in the mind of this author and had guided the objectives and the design of this investigation. The objectives include looking at the mode of evaluation, sample size, procedures, type of data collected to ascertain if the claims to effectiveness / superiority or otherwise being amplified in some quarters are valid or not.

Web-based learning was heralded with similar applause. While there are always rooms for improvement in all human endeavours, the initial evaluation results of WBL indicated, as described in Mioduser (2000); Inbal et al. (2003). (Current State of Websites in Science Education – Focus on Atomic Structure) “One step ahead for the technology, two steps back for the pedagogy”. They asserted that web technology has little to offer education in terms of pedagogy, science content, representation and communication. Almost all the 95 sites on atomic structure covered by Inbal et al. (2003) look the same as textbooks presentation. Although, many of them are student-centered, encouraging individualization but are lacking in inquiry and open-ended activities of the virtual environment online adaptive mechanism and thus encouraging memorization. In order to have objective evaluation as the one described above, there is the need to go beyond the affective referents and explore the cognitive and as well as pedagogical improvement brought about by the innovations.

Obviously, viewing the extent of interactivity, animation and excitement let alone the learning made possible through visualizations in the virtual lab, one will acclaim what a wonderful and high level technology has taken us.
Take the examples of Charles Sturt University (Dalgarno et al., 2003); Y Science Laboratories, Bingham at Young University, Virtual Chemistry Laboratory at Carnegie Mellon University, the use of tangible user interface in augmented chemistry in teaching organic chemistry (Morteu Fjeld and Botschi, 2007); Virtual Reality in Science and Engineering Education (Manseur, 2005); Discover Chemistry (Royal Society of Chemistry, 2009); Real-Time Visualization of the Quantum Mechanical Atomic Orbitals (R Dauger Research); Simulation software in a life sciences practical laboratory (Mcateer et al., 2002); A multi-user virtual-reality environment for a tele-operated laboratory (Hoyer et al., 2004) and several others. There is no doubting in the fact that, these efforts are profound as well as revealing. Yet, the pedagogical effectiveness needed to be proved beyond fun, enjoyment, attraction, motivation and other affective dispositions that dominate most of the times their evaluation reports. Comparative studies of educational goals and effectiveness of hands-on, simulated and remote laboratories by Corter et al. (2004); Cortet et al. (2007); Elawady and Tolba (2009); Taghavi and Colen (2009) yielded conflicting claims rather than gave a direction. It has also been reported that, many students take visualization literarily (Falvo, 2008; Kelly, 2005; Kelly and Jones, 2005; Srivinasan et al., 2006). Some perceived them as ‘fake’, ‘not authentic’ and therefore, rather than engendering understanding and enhancing performances, they tend to generate misconceptions and misrepresentations. Reports equally have it that, many students take visualization in science learning as fun (Falvo, 2008) and enjoyment (Stobart and Chau, 2002) the same way they take movies, computer games and other recreation modes. These conflicts can only be put to rest if there are enough data on pedagogical effectiveness to the two divides (These conflicts call for further inquiry into whether the position advanced by both advocates and critiques have enough data on pedagogical effectiveness).

**Research questions**

The research questions of interest in this study are:

(i) Are there enough evaluation reports beyond affective domain to justify the claim to pedagogical effectiveness of 3D environments and virtual realities over hands-on experience in real-laboratory?

(ii) What type of data was generated on 3D environments and virtual realities to back up the claim to pedagogical effectiveness?

The main hypothesis of concern in this study is that, there has not been enough evaluation to justify the claim to pedagogical effectiveness of technologically designed 3D realities over the hands-on experiences in the real-laboratory.

**METHOD**

The source of data in this study is literature search. The search was meant to locate literatures on evaluation of 3D world realities. The search and collection of data took eight months. This was carried out using the interactive portal, e-journal section of the University of Malaya library and through other online searches via the popular search engines such as http://yahoo.com; http://google.com; http://mamma.com; http://ask.com. The websites used in the university e-journal include Springer links, Informal world, ProQuest, Science Direct-Elsevier, Questia, IEEEExplore, Journals consulted include Computer and Education, International Journal of Science Education – all issues from 2000 - 2010 volumes, International Journal of Science and Technology Education – all issues from 2004 - 2009, Science Educators, Science Education, International Journal of Engineering Education, Chemistry Central Journal, Conference proceedings like those of SIGCSE 09 – March 2 - 7, 2009, SIGITE ’09 – October 22 - 24, 2009, WCCCE 09 – May 1 - 2, 2009, abstracts and theses downloaded from the internet. Articles located in all are over two thousands.

The researcher focused attention on topics that have to do with 3D environments, 3D world, virtual realities, virtual laboratory, visualizations, simulations, animations, and all other terms like remote laboratory, remote experiment, virtual experiments, hands-on laboratory, hands-on experiments etc in computer, engineering, architecture, science education, space science and so on. These were the links; the researcher looked for and also clicked on the search engines. To get nearer to the main objectives of the study the researcher needed to sort out articles that have to do with evaluation, effectiveness, assessment, potentials, performance, the role of and the benefits of. This narrowed down the sources to analyze. Next, the researcher examined the scope of the evaluation, the pedagogical implications, whether teaching or learning effectiveness, the learning outcomes covered, whether cognitive, psychomotor and affective, evaluation instruments, the sample size, the type of data generated, whether nominal, ordinal or intervals and the result of such evaluations. These were later analyzed to present results and draw conclusions.

**Limitation and significance of the methodology**

Whereas, the methodology involving secondary source data collection employed in this study is permitted, the authors want to appreciate the limitations of it. First hand data collection through experimental design or case study would, perhaps, yield better result, but the authors embarked on desk study for the purpose of getting enough data for better generalization. Another experimental design or case study would be too narrow. At the same time a survey design using questionnaire or interview would be impossible logistically and might not yield the type of data required in this study.

**Observations**

The search results indicated that the attentions on 3D world realities were diffused and diverse in literature. In a number of cases there were overlapping of ideas with other areas such as web-learning, problem-based learning, games and media. There were also sponsored projects which were not published in journals or proceedings but were uploaded as private web-sites on the internet. Examples of such could be found as sponsored projects by the universities, academic research institutes and agencies, and a
few individuals. One explanation for the diffused nature is that many of the sites, links and software are isolated and uncoordinated sporadic efforts. Many of them are business sites with only little educational implications. The domain that centers on education itself looks so few and relatively new ones with isolated efforts are particularly from the western world.

It was also discovered that articles and links that center on education were dispersed across disciplines; engineering and computer, natural sciences, education and psychology and so on with engineering and computer taking the lead. Those on natural sciences were divided among chemistry, physics, biology, space or earth science, environmental science, medicine, pharmacology with chemistry taking the lead. In education as a sole discipline, attentions were focused mainly on distance learning, supplemental software, self-paced interactivity and preparatory exercises ahead of real-lab experiments. The explanation for this trend was that the 3D environments are more of computer and engineering creation. Other disciplines of science and education just applied these computer technology and engineering to enhance their programs. Chemistry gets more attention because the discipline needs this application more for better comprehension of the micro and abstract nature of concepts that are prevalent in the study (Gabel, 1999; 2004; Johnstone, 2006). These concepts require modeling and visualizations to remove or reduce abstractions and for proper students' understanding (Dalgarno et al., 2003; Strangman et al., 2003). Distance education has gained prominence all over the world because of scarce education resources (Bayram, 1999; Mamoukari et al., 2000; Porter and Morgan, 2003). It is also an opportunity for those who could not make it to the four walls of the classroom to acquire diplomas and degrees. In order to give a touch of, or as a preparation for, real-lab hands-on experiences according to Bayram (1999); Mamoukari et al. (2000); Porter and Morgan (2003), many institutions have designed virtual-lab, visualizations, simulations and other 3D environments for students' learning in sciences out of school.

Another observation has to do with the issue of evaluation itself as a concept and in relation with the topic at hand. There were no standard criteria of evaluation of effectiveness due to what Ma and Nickerson (2006) describe as lack of agreement on conventions. While a few has an in-built evaluation procedures or models in their projects (Dalgarno and Headberg, 2001; Strangman et al., 2003) others made use of conventional evaluation models and instruments (Ealy, 1999; Morgil et al., 2008; Wang et al., 2007). Others still gave qualitative descriptions of potentials and benefits.

It was getting difficult to generate enough quantitative data for analysis. However, the author made use of the available articles (n = 24) with qualitative and quantitative data for the analysis.

RESULTS AND DISCUSSION

The sources of articles used and the number of them that dealt with evaluation are presented on Table 1 and Figure 1. The bulk of the articles consulted in this study as shown in Table 1 and Figure 1 are journal publications. The authors recognized that journal articles usually undergo better critical review than those in proceedings and non-affiliated ones. The reliance on journal articles as a source of secondary data collection was to add credibility to the study. The type of design, instrument, sample, data collected and result focus are presented on Table 2. The table above reveals that of the twenty four projects, 19 are based in the universities particularly from the western world (Dalgarno et al., 2009; Mayer, 2001; Schank et al., 2001; Stobart and Chau, 2002; Taghavi et al., 2009). Four can be classified as experiments (Barnea and Dori, 1999; Doris and Barak, 2001; Ealy, 1999; Thatcher, 2006) but in the actual fact, they were quasi-experiments because there were no randomization and control. Two of these projects combined both experiment and survey (Ealy, 1999; Taghavi and Colen, 2009) making use test and questionnaire as instruments for data collection. Qualitative data were mostly generated because questionnaire and interviews were more in use. Thirteen quantitative data were generated from tests, inventories, as well as from the questionnaires.

The research designs employed were mostly the case study and survey (83%), the instruments administered were questionnaire, interview and inventory (79%) and the type of data generated were largely qualitative (74%). As expected the findings revolved mainly around affective dispositions, and this can be seen in Figure 2.
Table 1. Classification of type of articles and their focus.

<table>
<thead>
<tr>
<th>Source of articles</th>
<th>Biology</th>
<th>Chemistry</th>
<th>Physics</th>
<th>Engineering</th>
<th>Computer</th>
<th>Medicine</th>
<th>Earth science</th>
<th>Architecture</th>
<th>Mathematics</th>
<th>No on evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journals</td>
<td>6</td>
<td>20</td>
<td>11</td>
<td>22</td>
<td>20</td>
<td>18</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Proceedings</td>
<td>2</td>
<td>10</td>
<td>7</td>
<td>12</td>
<td>18</td>
<td>2</td>
<td>3</td>
<td>-</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Uploaded articles</td>
<td>2</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>18</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Articles on evaluation</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 2. Evaluation parameters.

<table>
<thead>
<tr>
<th>Area of evaluation</th>
<th>Design</th>
<th>Instrument</th>
<th>Sample</th>
<th>Type of data</th>
<th>Learning outcome covered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experiment</td>
<td>Survey</td>
<td>Others</td>
<td>Test</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>University project</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>Agency/ass</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>5</td>
<td>14</td>
<td>-</td>
<td>6</td>
</tr>
</tbody>
</table>

Figure 2. Distribution of evaluation results across cognitive, skills and attitudes.

Although, those that examined concept formation and skill development showed enhanced performances and faster rates of learning, only two of the experiments (Barnea and Dori, 1999; Ealy, 1999) revealed significant difference between the treatment group and the non-experimental groups.

Apart from the positive qualitative expressions of satisfaction with benefits and potentials of these projects,
a few of the authors recounted negative comments or what look like areas for further improvement of the projects. Such comments include students not being clear of what is expected of them while manipulating the 3D structure of learning materials or in the virtual lab (Dalgarno et al., 2009; Ealy, 1999). Some of the students complained about waste of time and efforts, even when they did not have to measure, let alone get concerned about precision and accuracy as in the hands-on laboratories (Barak, 2007; Some expressed better understanding in 3D environments but preferred the hands-on laboratory to virtual laboratory (Corter et al., 2007; Fredman, 1997; Grabowski, 2007) because of the experiences there are exactly what they will be confronted with outside universities. Teachers complained of not being able to assist students because many of them lack the repertoire strategies of adventure in the 3D environments (Ealy, 1999). They also complained of lack of integration of these technologies with the real laboratory and classroom experiences.

Conclusion

Observation from this study reveals a great deal of technological innovations going on in all fields of human endeavours. The literatures consulted and used were dispersed across science disciplines. Many of them were not used because the domain of interest could not be attributed to education. Some could not be cited because they were not affiliated. Another important observation was that these efforts were mainly from universities and research institutes in American and Europe. Only sporadic efforts were reported from Asia and none from Africa. They were designed mainly with university laboratories. Teachers complained of not being able to assist students because many of them lack the repertoire strategies of adventure in the 3D environments (Ealy, 1999). They also complained of lack of integration of these technologies with the real laboratory and classroom experiences.

ACKNOWLEDGEMENT

We acknowledge the Management of University Malaya for the use of the interactive portal of the university library where enormous data were gathered for this study.

REFERENCES


