Experience with Virtual Reality-Based Technology in Teaching Restorative Dental Procedures

Judith A. Buchanan, Ph.D., D.M.D.

Abstract: This article reports on extensive experience with advanced simulation at the University of Pennsylvania, School of Dental Medicine (UPSDM). Virtual reality-based technology (VRBT) or advanced simulation is currently available for the instruction of dental students in preclinical restorative procedures. UPSDM was one of the first schools in the world to have extensive experience with VRBT technology using an advanced simulation unit (DentSim) from DenX, Ltd. UPSDM’s experience consists of several years of research using control and experimental groups, employing students to participate in an investigative project, and using the units for remediation and a supplement to the preclinical laboratory. Currently, all first-year students (Class of 2007 and Class of 2008) are receiving most of their preparative operative training on the VRBT units. UPSDM started with one (beta) version unit in 1998, which was later updated and expanded first to four units and recently to fifteen units. This communication is presenting the studies that were of fundamental importance in making the decision to acquire fifteen units in 2003. The areas of main interest to the SDM concerning this technology were its use in teaching, refreshing, and remediating students in restorative procedures and its effectiveness as a teaching methodology in relation to time, efficiency, and faculty. Several studies with varying parameters were performed at various time points. The limited statistical analysis conducted was not conclusive for all measures, and in some cases the data only suggest areas of possible significance. This is due to the low number of students who could access the limited number of available units and the change of protocols in response to student and faculty input. Overall, the results do suggest, however, that students learn faster, arrive at the same level of performance, accomplish more practice procedures per hour, and request more evaluations per procedure or per hour than in our traditional laboratories. Students’ attitudes, as measured by surveys, group interviews, and private interviews, are mixed. Our overall evaluation of our experience with this technology has been positive and led to the purchase of additional units, its full incorporation into our curriculum, and curriculum revision to maximize its potential. Our conclusion is that this technology offers significant potential in the field of dental education and that further use and investigation are both desired and justified.

Dr. Buchanan is Associate Dean, Academic Affairs, Department of Preventive and Restorative Sciences, University of Pennsylvania, School of Dental Medicine. Direct correspondence and requests for reprints to Dr. Buchanan at the Office of Academic Affairs, University of Pennsylvania School of Dental Medicine, 240 S. 40th Street, Philadelphia, PA 19104-6003; 215-898-1167 phone; 215-573-4761 fax; Jbuchana@pobox.upenn.edu.

This project was funded by a grant from the University of Pennsylvania Research Foundation.

Key words: virtual reality, dental restorative procedures, dental education, technology, computer-assisted instruction

Submitted for publication 8/12/04; accepted 10/19/04
of restorative preparative procedures. These potential advantages include the ability of the unit to give consistent, unbiased feedback based on evaluation of the preparation in terms of tenths of millimeters. The instant availability of this feedback and the unlimited availability of the unit can also be viewed as potential advantages. Basically, the unit is composed of “real” items such as a simulated patient head and torso, KaVo dentoform, handpieces, air/water syringe, dental light, and an imaging system that translates the operator’s tooth preparation into a virtual image within a computer and displays it on a screen.

Using the VRBT unit, a student can prepare dentoform teeth in much the same manner as with the common dental simulators available at many schools. However, the unique property of this unit is its ability to construct a real-time virtual image of the student’s preparation in the computer. The computer can evaluate the tooth preparation both immediately and at the student’s request. Real-time evaluation for critical, noncorrectable errors is given as immediate feedback while more complete, detailed evaluation of a restorative preparation is given when requested by the student. The extensive evaluative feedback given when requested is presented in visual and written forms and includes a numerical grade. Hence, the VRBT unit offers objective, consistent evaluation of preparations easily obtained at any time during the process of preparing the tooth. This evaluation includes both formative (corrective feedback) and summative (resulting in a final grade) evaluation. This is in contrast to an evaluation given by faculty that consists, for the most part, of evaluation of an end product.

Although virtual reality-based or advanced technology is available in dental education, there is limited research investigating the effectiveness of this technology. The available research has produced conflicting results and needs further review.\textsuperscript{22-27} VRBT has matured considerably since its introduction, and it is important to take into consideration the differences of experience between schools and the level of maturity of the unit that was used in different studies when reviewing the outcomes of VRBT research studies.

The University of Pennsylvania School of Dental Medicine (UPSDM) was one of the first dental schools in the world to purchase a VRBT unit and the first to incorporate this technology into its curriculum in 1998. Over the next five years, UPSDM investigated the effectiveness of this technology in a variety of ways and, in the late fall of 2002, made the decision to not only purchase a significant number of units, but to modify our curriculum to take advantage of the potential of VRBT. This article describes our experience with the technology that led to this decision and reviews the results of three studies conducted at UPSDM.

### Description of VRBT Studies at UPSDM

Our first study began soon after we had acquired a VRBT DentSim unit. Being one of the very first units, it could be classified as a beta unit. Although it was much more prone to technical problems than later units, the important and unique features were consistent with the state of the technology at that time. A sample of sixteen students was selected from the ninety-four students enrolled in the first-year class in the following manner. Students were first grouped based on three measures: learning style/type combination, psychomotor skill level, and the Dental Admission Test (DAT) perceptual score. Learning styles were determined on all incoming students using both the 4Mat and Kolb learning style inventories.\textsuperscript{28,29} After students were grouped, student pairs were selected within each group by matching as closely as possible for the remaining sample characteristics, which were computer comfort (determined by survey), gender, DAT academic averages, dental school academic averages, and college academic averages. For each of the eight pairs, one student of the pair was randomly chosen for the experimental group and the other assigned to the control group. This assignment ensured that the control group and experimental group were very similar in composition, a factor especially important with small groups. The sample size was limited to sixteen students based on the maximum projected use of one VRBT unit by eight students in the experimental group during the time allotted for the restorative procedure within the existing preclinical operative course. The study sample consisted of twelve men and four women.

Experimental students were assigned to the unit in two-hour time blocks over a three-week time period. Students in the control group were instructed in the UPSDM traditional laboratory. Each station in the traditional laboratory was equipped with a light, handpiece connections, and a post with an attached Dentoform. Manikin heads were not routinely used,
and it should be noted that the variable of more or less sophisticated physical simulation may affect attitudes toward the methodology. The experimental group was instructed to practice Class I amalgam procedures on the simulation unit and to receive no instruction on this procedure from faculty, other students, or monitors. The control group prepared teeth in the traditional lab and requested evaluation from faculty.

Although the control group had access to faculty only during course hours, they were allowed to practice during the hours the laboratory was open (nights and weekends). In both groups, the criteria for assessment were determined by the restorative faculty involved in the preclinical operative course and were consistent between the two groups. Both groups were instructed to keep detailed time logs in which they recorded 1) time used to practice preparations, 2) the number of teeth used for practice, and 3) number of times they requested expert evaluation. The number of times students asked for expert evaluation represents the number of times an experimental student asked the computer for a complete evaluation of their preparation or when a control student asked a faculty member for evaluation. Weekly meetings and post-experiment personal interviews with each student affirmed that students had made every effort to keep accurate logs, although I acknowledge the possibility that some inaccuracy may be inherent in data that is self-reported.

At a scheduled time, a practical examination (Class I preparation on tooth number 19) was given to both control and experimental groups. The teeth were collected from both groups, and one faculty member graded all teeth from both groups in a blinded manner.

The second study was initiated after UPSDM obtained four VRBT units. Along with the addition of more units, the software of the units had been updated several levels. The format from the first study was repeated except for the following items. Twenty-eight students were chosen and paired only for the following parameters: DAT perceptual score, learning style/type combination, gender, DAT academic averages, and dental school academic averages. The determination was made randomly as to which student of the pair was assigned to the control versus experimental group, with fourteen students in each group. Students learned the procedures of Class I and Class II amalgams, and limited faculty input was allowed. Faculty input was available for each student one hour a week while they were assigned to the VRBT laboratory. Faculty only responded to questions such as what burs to use, how to correct errors, when to use the slow speed, etc. The faculty did not evaluate those parameters evaluated by the unit that used the same criteria as the faculty used to evaluate the control group. Detailed logs were kept in an identical manner as in Study 1. Again, student performance was measured by a practical examination of a Class I and Class II preparation, which was graded by one faculty member in a blinded manner.

Although the technical problems, especially with calibration, improved in each upgrade of the software, there was frustration among students who participated in the experimental groups of the previous studies when technical problems interfered with their ability to complete an assignment that was part of their course grade. To remove the pressure of having the experience as part of their course, it was decided that further study of the VRBT would be done outside of course requirements. In Study 3, seven dental students between their freshman and sophomore years were employed by the school and worked in the Virtual Reality Laboratory five hours per day for seventeen days. A dentist was present in the room at all times to help with the technology and answer questions that were not related to preparation evaluation. Students reached competency in amalgam Class I, Class II, and Class V and composite Class III, Class IV, and Class V preparations. The students also worked on crown preparations (full gold, porcelain fused to metal, all porcelain). Students kept logs of the number of practice preparations and times they asked for evaluation per preparation.

One important aspect of the incorporation of new technologies into the curriculum or into dental practice is the level of acceptance by users. Hence, students’ attitude toward and acceptance of this technology were evaluated by surveys and personal interviews for all three studies.

**Results and Discussion**

The first study included eight students each in the control and experimental groups. To verify that control and experimental students had similar levels of psychomotor skills, we compared the two groups’ overall laboratory grade (didactic coursework grade is separate from the laboratory grade) in the operative course and found no statistical difference (p=0.50). This grade was a summation of nine practical exams. Control students averaged 6.71 hours pre-
paring Class I preparations compared to the experimental students who spent an average of 3.69 hours (Table 1, p=0.07). Experimental students asked for evaluation more times (6.52 evaluation per hour of practice) than control students (3.21 evaluations per hour of practice) (Table 1, p=0.08). The experimental students prepared significantly more teeth per hour (average of 3.8 preparations/hr) than the control students (average of 1.6 preparations/hr) and not surprisingly used more teeth (average of 11.71 versus 6.57 for control, p=0.02). This emphasizes a higher efficiency in preparing teeth using VRBT. The results of the Class I practical examination were slightly higher for the experimental group (72.6 percent versus 61.9 percent) but were not statistically significant (p=0.23, Table 2).

In Study 2, which contained fourteen students in each group (control and experimental), similar results were found with control students spending 14.2 hours practicing Class I and II preparations before the practical as compared to the experimental group spending 8.2 hours practicing (Table 1). Control students in Study 2 prepared 2.1 preparations per hour compared to the experimental group, which prepared 3.4 preparations per hour. The control group requested an average of 1.99 evaluations per hour of practice compared to the experimental group, which requested 6.18 evaluations per hour (Table 1). The most obvious reason for this difference may be the ease of requesting evaluations from the computer compared with the difficulty or reluctance of a student to attract the attention of a faculty member for evaluation. These evaluations were both formative and summative. The fact, however, that the students requested evaluations this frequently suggests that the students in this study viewed frequent feedback as helpful to learning. This frequency of evaluations would not be feasible in a traditional laboratory since it would require a faculty member supervising twelve students to give a complete evaluation of a preparation about every forty-five seconds.

Recent experience with VRBT using an entire class has shown that students need to be weaned from the availability of instant evaluation at some point in order to develop the ability to self-assess using instruments always available such as a probe and condenser.

In Study 2, students in both control and experimental groups completed two practicals that were graded by one faculty member in a blinded manner. Control students scored 79.3 percent and 80.9 percent on the Class I and Class II practicals respectively, while the experimental students scored lower than control with 72.0 percent in the Class I and similar to control in the more difficult Class II at 79.7 percent (Table 2). The two groups’ final overall grades in the course (laboratory portion only) were very similar, with the control students earning 86.0 and the experimental groups averaging 86.7.

---

### Table 1. Comparison of time and evaluations requested

<table>
<thead>
<tr>
<th>Groups</th>
<th>Time Spent Practicing Per Hour of Practice</th>
<th>Evaluations Requested Per Hour of Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Study 1</td>
<td>6.71 hrs (Class I)</td>
<td>3.21</td>
</tr>
<tr>
<td>Experimental Study 1</td>
<td>3.69 hrs (Class I)</td>
<td>6.52</td>
</tr>
<tr>
<td>Control Study 2</td>
<td>14.2 hrs (Class I &amp; II)</td>
<td>1.99</td>
</tr>
<tr>
<td>Experimental Study 2</td>
<td>8.2 hrs (Class I &amp; II)</td>
<td>6.18</td>
</tr>
<tr>
<td><strong>Combined groups (Studies 1 &amp; 2)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control groups’ averages</td>
<td>20.1 hrs</td>
<td>2.41</td>
</tr>
<tr>
<td>Experimental groups’ averages</td>
<td>11.8 hrs</td>
<td>6.29</td>
</tr>
</tbody>
</table>

Students recorded in log books the hours they spent practicing before the practical exams and the number of times they requested expert evaluation. For control students, this was the number of times they requested feedback from an instructor; for experimental students, it was the number of times they requested evaluation from the unit. Study 1 had eight students per group; Study 2 had fourteen students per group.

### Table 2. Comparison of practical examinations and final course grades

<table>
<thead>
<tr>
<th>Group</th>
<th>Practical Examinations</th>
<th>Final Course Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class I</td>
<td>Class II</td>
</tr>
<tr>
<td>Control Study 1</td>
<td>61.9</td>
<td>89.3</td>
</tr>
<tr>
<td>Experimental Study 1</td>
<td>72.6</td>
<td>87.8</td>
</tr>
<tr>
<td>Control Study 2</td>
<td>79.3</td>
<td>80.9</td>
</tr>
<tr>
<td>Experimental Study 2</td>
<td>72.0</td>
<td>86.0</td>
</tr>
<tr>
<td><strong>Combined averages (Studies 1 &amp; 2)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control groups’ averages</td>
<td>70.6</td>
<td>80.9</td>
</tr>
<tr>
<td>Experimental groups’ averages</td>
<td>72.3</td>
<td>87.7</td>
</tr>
</tbody>
</table>

Results of practical examinations of procedures evaluated by one faculty member in a blinded manner. Study 1 involved eight students per group; Study 2 involved fourteen students per group.
It should be noted that there are separate grades for the psychomotor portion and the didactic portion of this course and only the psychomotor portion was used for comparison.

Data from the third study, in which seven dental students were employed to work on the VRBT unit, shows that these students requested evaluations at a frequency seen in earlier studies (average of 7.60 evaluations per procedure). Students prepared an average of 4.4 crown preparations per four-hour session. The crown preparations were impressive to our faculty even though these students had no previous didactic or laboratory experience with crown preparations. Since these students had considerable experience with both the VRBT (over eighty-four hours total) and traditional laboratory (approximately 300 hours), they were asked to rate on a 1 to 5 scale the VRBT training as compared to their experience with the traditional laboratory. Results are shown in Table 3. All students fully agreed that the VRBT unit improves manual skills, improves minor movements with a handpiece, and increases speed. The lowest agreement was in the statement that the VRBT widens the theoretical basis of cavity preparations (3.71 out of 5) and that it gives accurate evaluations (3.29 out of 5). In Figure 1, responses of students directly comparing the traditional laboratory (LAB) to the VRBT in learning tooth preparation are shown. Students were asked if VRBT was more helpful, less helpful, or the same as the traditional lab for learning. Students were positive about some aspects of VRBT as compared to LAB, with the majority indicating that VRBT provided more feedback, allowed self-evaluation, and was more like a real patient situation. It must be noted, however, that the more sophisticated manikin in the VRBT unit, by itself, could influence this response. Students felt the VRBT was harder to use than the traditional setup in the preclinical laboratory. This could be attributed to technical difficulties with earlier units and software versions.

Additional surveys of students in all three studies (data not presented) and focus group discussions completed by all students involved with the VRBT over several years’ time span can be summarized into several key observations: 1) students want faculty to play some part in their psychomotor training even though students dislike the considerable variability between instructors and are frustrated by waiting for faculty; 2) students view VRBT technology as having a positive role in preclinical training; 3) students feel that they learn faster with VRBT; and 4) students feel more confident with a high-speed handpiece after training on VRBT.

Although several other schools have similar positive experiences, there are some who have had experience with VRBT and have drawn different conclusions about its potential. A comparison of these other experiences is necessary to help put the different opinions in perspective. Quinn et al. reported on the Dublin Dental School and Hospital’s experience with VRBT. This institution’s pilot project, which used very early generations of simulators, measured student performance after only four hours of exposure to VRBT plus sixteen hours of traditional teaching and, not surprisingly, showed no difference in student performance between a control group (twenty-one hours of traditional training) and experimental groups. A study published soon after led the same authors to make the statement that “VR-based skills acquisition is unsuitable for use as the sole method of feedback and evaluation for novice students—a very strong statement indeed. Further inspection of the article raises concern over the support for this statement.

The data presented in the second study by Quinn et al., which again used early generation units, is derived from twenty students randomly divided into two groups. The control group used the VRBT unit for five hours but rather than using the evaluation capabilities of the unit, relied on faculty feedback. The experimental group also used the VRBT unit for five hours but received all evaluation and feedback from the unit. After the practice ses-

<table>
<thead>
<tr>
<th>Table 3. Summary of student evaluation of VRBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions Rating VRBT (after over 80 hours’ experience)</td>
</tr>
<tr>
<td>User friendly</td>
</tr>
<tr>
<td>Improves manual skills</td>
</tr>
<tr>
<td>Improves minor movements</td>
</tr>
<tr>
<td>Increases speed</td>
</tr>
<tr>
<td>Widens theoretical basis</td>
</tr>
<tr>
<td>Assists in 3D visualization</td>
</tr>
<tr>
<td>Improves self-assurance</td>
</tr>
<tr>
<td>Major teaching source</td>
</tr>
<tr>
<td>Provides detailed feedback</td>
</tr>
<tr>
<td>Accurate evaluation</td>
</tr>
</tbody>
</table>

Seven students evaluated their experience with VRBT after 80 hours’ experience. This experience included procedures they had learned previously in a traditional laboratory and procedures they had neither previous experience with nor didactic foundational lectures.
sions, both groups executed two Class I cavity preparations that were graded in a blinded manner by two faculty scorers. The authors reported there were statistical differences between the control and experimental groups in three out of five criteria for assessment, with the experimental group performing less well. This data is suspect since there was no effort to match control or experimental groups in any parameter such as innate psychomotor ability, academic performance, gender, previous computer experience, etc. In fact, no data concerning the makeup of the control versus experimental students was provided, including their overall performance in preclinical courses. As many operative faculty will attest to, students can vary widely in their innate ability to grasp the psychomotor skills of operative dentistry, and at the very least, the control group versus experimental group performance in other psychomotor courses or other preparations in the same course should have been compared. It is possible that this situation could have been repeated between two randomized groups of ten students’ learning one operative procedure in completely identical ways. If randomized selection was used, it would have been wise to have two control groups to strengthen any possible conclusions. If there was no statistical difference between the two control groups on the evaluation of a preparation in any of the five criteria, differences of two or three of the five criteria between the two control groups and the VRBT experimental group may have been more valid. To use the data as presented in that publication to nullify a technology and state that this method of teaching is unsuitable seems to be far from justified. Regardless of the design, it would appear unwise to most in dental education to make such a strong statement based on five hours of experience of ten students. Our experience with the VRBT spanning over five years and involving several generations of hardware and software, multiple studies, countless student hours, numerous surveys and student discussions, and many restorative procedures gives us the confidence that our results are valid and the decision to further embrace this technology is justified.

Figure 1. Student summary evaluations comparing VRBT to traditional laboratory

Students from Study 3 evaluated their experience with VRBT after several months’ experience with the traditional laboratory. They were asked to compare VRBT with the traditional laboratory as to whether the VRBT was more helpful, the same, or less helpful. Graph is courtesy of Dr. Eldad Irani.
Conclusion

Although our goal was to scientifically demonstrate the efficiency of VRBT in the education of psychomotor skills, we soon realized that this was a very difficult task due to various factors and limitations during late 1990s when these units were first available. The limitations of low numbers of units and hence low numbers of students, technical problems, the desire to modify the educational environment as soon as we realized a better way to do it, and the frequent upgrades of software and hardware did not lend itself to the production of statistically significant data. Most of our data is suggestive, but is supportive of VRBT as an educational tool, and some data demonstrated statistically significant differences between experimental and control groups. But as Dr. Kathleen Rosen20 remarks in a recent book devoted to health care simulation, “no one has ever proved the validity or value of simulation education and training in any other industry” before its accepted use. Hence, it may be that we cannot wait to adapt VRBT simulation until unequivocal proof of the benefit had been demonstrated.

In addition to the described research data we obtained, we formed opinions of the value of this technology in many other, nonmeasurable ways. Considerable time was spent watching students in the VRBT lab, talking with them at length, listening to them talk to the many visitors who came to observe, and watching students demonstrate the VRBT to others at our school and at national meetings. The sum of all our data, experience, evaluation, observations, and intuition resulted in the belief that VRBT is a powerful educational methodology with the potential to significantly affect dental education. This belief was the basis for the purchase of fifteen VRBT units, allowing for the full incorporation of this technology into our courses and associated curriculum revision to help this technology realize its full potential in dental education.

Acknowledgments

I would like to thank Dr. Denise Stewart, Dr. Joan Gluch, Dr. Amer Abu-Hanna, Dr. Penny Hammrich, Dr. Francis Mante, Dr. Mamle Mante, and Dr. Eldad Irani for their help with these studies and Dr. Peter Berthold for help with the preparation of this manuscript.

REFERENCES


24. Abstract by Case Western Reserve University, Cleveland, OH.


