Effectiveness of Web-Based Teaching Modules: Test-Enhanced Learning in Dental Education


Abstract: The purpose of our study was to evaluate the effectiveness of self-tests as a component of web-based self-instruction in predoctoral orthodontics and pediatric dentistry. To this end, the usage patterns of online teaching modules and self-tests by students enrolled in three courses at the University of North Carolina at Chapel Hill School of Dentistry were monitored and correlated to final exam grade and course average. We recorded the frequency of access to thirty relevant teaching modules and twenty-nine relevant self-tests for 157 second- and third-year D.D.S. students during the course of our data collection. There was a statistically significant positive correlation between frequency of accessing self-tests and course performance in one course that was totally based on self-instruction with seminars and multiple-choice examination (Level IV): Spearman correlation between frequency of self-test access and final exam grade, rho=0.23, p=0.044; correlation between frequency of self-test access and course average: rho=0.39, p=0.0004. In the other two courses we monitored, which included content beyond self-instruction with self-tests, the correlations were positive but not statistically significant. The students’ use of online learning resources varied significantly from one course (Level I) to the next (Level II): Wilcoxon matched pairs signed-rank tests, S=−515.5, p=0.0057 and S=1086, p<0.0001. The data from this study suggest that increased use of web-based self-tests may be correlated with more effective learning in predoctoral dental education by virtue of the testing effect and that dental students’ usage of resources for learning changes significantly over the course of their education.

Keywords: dental education, computer-assisted learning, computer-based instruction, test-enhanced learning, orthodontics education, pediatric dentistry education

Submitted for publication 5/12/10; accepted 8/26/10
required when compared to more traditional dental school lecture courses.

A systematic review of the literature assessing the effectiveness of computer-aided self-instructional programs in dental education has shown that this type of self-directed and technology-assisted learning is at least as effective, and in many cases more effective, than more traditional modalities of learning. It has also been demonstrated that students perceive computer-assisted self-instruction as an effective and acceptable way to learn. We know that students accept the use of these teaching modules and self-tests and that they learn well with this approach. We do not yet objectively know why this method is successful or how to optimize students’ use of it. In this study, we evaluated the effectiveness of self-tests as a component of web-based self-instruction in predoctoral orthodontics/pediatric dentistry by monitoring and analyzing the usage patterns of online self-tests by students enrolled in three courses.

Methods

Participants in our study were the members of the D.D.S. classes of 2011 and 2012. The class of 2012 (N=78) was enrolled in the Level I Growth and Development course during the summer trimester of 2009 and in the Level II course during the fall trimester of 2009. The class of 2011 (N=79) was enrolled in the Level IV Growth and Development course during the fall trimester of 2009.

During each section of each course, students viewed a group of online teaching modules and then participated in a small-group seminar for further discussion of the material in which their performance was graded. All students enrolled at the UNC School of Dentistry with a valid login have access to these web-based teaching materials at any time. The school’s Office of Computing and Information Systems (OCIS) monitored student access to both the teaching modules and the self-tests, each of which is contained in a separate Flash program that has limited internal tracking capabilities as constructed. Accordingly, OCIS was able to record the date and time of access to each individual module or self-test but not the duration of use or any details that might indicate the validity of each access.

For the Level I course, a student’s course average was determined based on their performance on the free-response tests after each of the four small-group seminars (70 percent) and one multiple-choice final exam (30 percent). In the Level II course, the overall average was determined by two in-class free-response exams (30 percent), a multiple-choice and free-response final exam (50 percent), and a series of projects including seminar participation and exercises related to the course material (20 percent). The Level IV course average was determined based on one in-class multiple-choice final examination (80 percent) and participation in course seminars (20 percent). Each student was assigned a dummy identifier, and once the access records and grades were compiled, his or her access data were associated with both the final exam grade and course average by the Orthodontics Department manager, who neither directed any of these courses nor participated in data analysis for this study.

Summary statistics were computed for the access records as well as the final exam grades and course averages. Frequency of access of modules and self-tests was correlated separately to both final exam grade and course average for each participant in the Level I, II, and IV courses using Spearman’s correlation.

The change in frequency of access of modules and self-tests for each participant from Level I to Level II was analyzed using a Wilcoxon matched pairs signed-rank test. Only data from the seventy-six participants who were enrolled in both Level I and Level II were considered in this test. The change in frequency of access of modules and self-tests for each participant from Level I to Level II was correlated to the change in final exam grade and course average for each participant from Level I to Level II using Spearman’s correlation. For all tests, statistical significance was set at p=0.05, and SAS Version 9.1 was used for data analyses.

Results

Summary statistics for data collected over all three courses appear in Table 1. The data did not follow a normal distribution for any of the variables considered.

For Levels I and II, Spearman’s correlations between frequency of accessing modules and self-tests and final exam grade and course average indicated that there was no significant relationship (p>0.05 for all correlations) between the frequency of access to either the teaching modules or the self-tests or to performance in the course as indicated by final exam grade or course average. For the Level IV
course, there was again no statistically significant correlation between module access (Figure 1) and either final exam grade or course average (p>0.05 for both). There was, however, a statistically significant relationship between frequency of access to the Level IV self-tests and both the final exam grade (rho=0.23, p=0.04) and the course average (rho=0.39, p=0.0004) (Figure 2). When possible outlier data (self-test ac-

Table 1. Summary statistics for data collected over the 2009 Level I, II, and IV courses

<table>
<thead>
<tr>
<th>Course</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Range</th>
<th>Interquartile Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module Access Frequency†</td>
<td>Level I 78</td>
<td>41.4</td>
<td>38.0</td>
<td>18.6</td>
<td>10–104</td>
<td>27.5–51</td>
</tr>
<tr>
<td></td>
<td>Level II 78</td>
<td>32.8</td>
<td>31.0</td>
<td>19.0</td>
<td>2–103</td>
<td>21–42.3</td>
</tr>
<tr>
<td></td>
<td>Level IV 79</td>
<td>18.4</td>
<td>17.0</td>
<td>11.5</td>
<td>0–81</td>
<td>11–23</td>
</tr>
<tr>
<td>Self-Test Access Frequency†</td>
<td>Level I 78</td>
<td>34.0</td>
<td>30.5</td>
<td>23.3</td>
<td>0–103</td>
<td>15.5–51.3</td>
</tr>
<tr>
<td></td>
<td>Level II 78</td>
<td>12.1</td>
<td>11.0</td>
<td>9.6</td>
<td>0–39</td>
<td>4.3–18</td>
</tr>
<tr>
<td></td>
<td>Level IV 79</td>
<td>19.9</td>
<td>17.0</td>
<td>14.5</td>
<td>0–78</td>
<td>9.5–27</td>
</tr>
<tr>
<td>Final Exam Percentage</td>
<td>Level I 78</td>
<td>82.2</td>
<td>82.0</td>
<td>6.9</td>
<td>64–98</td>
<td>78–86</td>
</tr>
<tr>
<td></td>
<td>Level II 78</td>
<td>83.4</td>
<td>84.2</td>
<td>6.9</td>
<td>68–95.1</td>
<td>79–88.5</td>
</tr>
<tr>
<td></td>
<td>Level IV 79</td>
<td>96.8</td>
<td>100.0</td>
<td>4.7</td>
<td>80–100</td>
<td>96–100</td>
</tr>
<tr>
<td>Course Average Percentage</td>
<td>Level I 78</td>
<td>91.0</td>
<td>90.7</td>
<td>2.9</td>
<td>85.9–99.1</td>
<td>88.7–93.1</td>
</tr>
<tr>
<td></td>
<td>Level II 78</td>
<td>89.4</td>
<td>89.9</td>
<td>4.1</td>
<td>81.1–96.5</td>
<td>87–93</td>
</tr>
<tr>
<td></td>
<td>Level IV 79</td>
<td>95.4</td>
<td>96.8</td>
<td>4.0</td>
<td>83.6–100</td>
<td>93.4–98.4</td>
</tr>
</tbody>
</table>

†Students had access to twelve modules and self-tests for Level I, eight modules and self-tests for Level II, and ten modules and nine self-tests for Level IV.

Figure 1. Spearman’s correlations between frequency of access of modules and final exam (rho= -0.10, p=0.37) and course average (rho= -0.046, p=0.68)
cess frequency of 78, module access frequency of 81) for one participant in the Level IV course were excluded from the data set, only the Spearman’s correlation of self-test access frequency to course average was still found to be statistically significant (rho=0.38, p=0.0006).

Wilcoxon matched pairs signed-rank tests showing change in frequency of access of modules and self-tests from Level I to Level II are shown in Figures 3 and 4. This shows that participants accessed teaching modules more frequently in Level II than Level I (S=−515.5, p=0.006) and that they accessed self-tests less frequently in Level II than Level I (S=1086, p<0.0001). Spearman’s correlations between change in frequency of access of modules and of self-tests and change in final exam grade and course average from Level I to Level II revealed no statistically significant relationship between any of these variables (p>0.05 for all).

Discussion

For more than a century, cognitive psychology research has shown that learning is strongly enhanced by the repeated spaced presentation of material. More recent investigations have demonstrated that frequent testing has the powerful ability to enhance learning itself by what is called “the testing effect”: more is gained by testing while being exposed to material than simply by studying and taking a final exam. It has even been shown that repeated testing enhances learning more than studying or restudying material; there is a benefit in testing beyond simply being exposed again to the same material in a different way. Finally, after testing, feedback increases the capacity for learning, especially when multiple-choice tests with distracters are used. While much of the psychology research on the testing effect up to this point has been conducted in laboratory settings or in university classroom environments, it has been reported that exposure to online self-assessment tools can improve course performance. And although it has been suggested that the principles of test-enhanced learning might be used effectively to enhance predoctoral medical education, we do not yet have strong evidence that computer-assisted self-instruction in a dental school setting can be optimized by applying these ideas.
Figure 3. Wilcoxon matched pairs signed-rank test for change in average module access from Level I to Level II (Q1 change=-1.71, median change=-0.67, Q3 change=0.58, S=-515.5, p=.0057), demonstrates that students accessed teaching modules more in Level II than in Level I in a statistically significant way.

Figure 4. Wilcoxon matched pairs signed-rank test for change in average self-test access from Level I to Level II (Q1 change=0.15, median change=1.19, Q3 change=2.35, S=1086, p<0.0001), demonstrating that students accessed self-tests less in Level II than in Level I in a statistically significant way.
The self-test component of the Growth and Development curriculum offers a vehicle to employ the psychological principles of test-enhanced learning in a predoctoral dental education setting. The self-tests are designed so that the student receives immediate feedback after answering each multiple-choice question and each self-test can be accessed multiple times and at the student’s convenience. Despite the fact that research has shown that scores on low stakes tests, such as the Growth and Development self-tests, tend to be slightly lower than high-stakes tests, low-stakes tests are as reliable.\(^\text{18}\) If we are to assume that the testing effect is a valid pedagogical tenet for predoctoral dental education, then it would be logical to expect that students who use these online self-tests more frequently would have a more effective learning experience as indicated by better performance on course examinations.

**Class of 2011**

The expected result that improved course performance correlates with more frequent self-test access is indeed what we saw for the Class of 2011 in the Level IV course. Simply looking at the module information more frequently did not appear to enhance learning, while taking more self-tests did (Figures 1 and 2). It seems that the testing effect optimized learning for the Class of 2011 in the Level IV course, which is taken in the third year of dental school as the terminal Growth and Development course. The structure itself of the Level IV course may have mitigated confounding factors in our evaluation of the effects of study resources on learning in an undisturbed educational setting. The Level IV course uses only one multiple-choice examination for evaluation and employs a bank of test questions developed directly from the Level IV teaching modules and self-tests. The final exam questions are not, however, drawn from these sources in such a way that a student’s performance on the final exam could be a reflection of simply having seen the actual exam questions previously.

It must be acknowledged that the correlations between self-test access and final exam score or course average are weak and that the correlation to final exam score may not be robust. With a possible outlier data point eliminated, correlation of self-test access frequency to course average remained statistically significant while correlation to final exam score did not. It is not clear that this data point is indeed an outlier since the data does not follow a normal distribution and since the participant who accessed the self-tests seventy-eight times also accessed the teaching modules eighty-one times. It is interesting to note that because 80 percent of the course average for Level IV is determined by the final exam grade, the course average is in fact primarily a reflection of final exam performance.

**Class of 2012**

For both the Level I and Level II courses, there was no statistically significant positive or negative correlation between more frequent access of modules or of self-tests and improved course performance. Why might these results differ from those obtained in the Level IV course? A full year more of dental education may have influenced the study habits of the Class of 2011, leading them to a different pattern of activity. Differences in course structure and study resources available may offer a more relevant explanation of these results, however.

Level I Growth and Development is the introductory course of the series and is taken over the summer trimester in the first year of dental school. While the final exam is a multiple-choice examination with questions developed almost exclusively from the teaching modules and self-tests, the four other exams in the course (at the end of each unit, after the small-group seminar) are in a free-response format in which students must produce written short answers. This style of testing—production testing—has been shown to promote better retention of material than recognition (i.e., multiple-choice) tests.\(^\text{5,19}\) In this way, this course includes more examinations than either Level II or IV and provides an additional effective learning tool by virtue of these examinations. Additionally, many students have access to test files with the unit exams from previous years that serve as study aids. Since these exams are not returned to the student’s possession after grading for this purpose, students are unlikely to report use of this study tool. These extra free-response examinations and undocumented test-file exams provide confounders to investigating the effectiveness of the Level I self-tests.

Level II is the second course of the series, taken in the fall trimester of the second year of dental school. Importantly, this course covers material presented in lectures and clinical experiences in pediatric dentistry as well as material in the online teaching modules. The course includes two free-response examinations and a final exam as well as a series of projects that contribute to the overall course...
average. Forty percent of the material covered on the final exam is not covered in the Growth and Development online materials and contributes 50 percent of the course average. The course structure, presence of free-response exams, and grading formula all clearly have the potential to introduce confounding forces into the correlation of module and self-test usage with final exam grades and course averages.

Changes from Level I to Level II

The Class of 2012 maintained seventy-six of its seventy-eight members across the Level I and Level II courses. A Wilcoxon matched pairs signed-rank test shows that they accessed modules more frequently (Figure 3) and self-tests less frequently (Figure 4) in Level II than in Level I. While neither the respective increase nor decrease can be correlated to any meaningful change in final exam score or course average, it is reasonable that the same confounders of course structure and study resources mentioned above would be influential here. The lack of correlation in changes from Level I to Level II is not as interesting as the fact that there is clear evidence that the participants’ use of learning resources changed significantly from one course to the next. It is substantiation that dental students’ study methods may evolve appreciably over the course of their education.

Conclusions

This study suggests that increased use of web-based self-tests may be correlated with more effective learning in predoctoral dental education by virtue of the testing effect and that dental students’ usage of resources for learning changes significantly over the course of their education. These are considerations relevant for the many dental schools that face shortages of full-time faculty to teach a congested curriculum and are seeking innovations in both their educational programs and how they are taught.

Acknowledgments

We thank Dr. Ceib Phillips for statistical consultation, Dr. Aimee Callender and Dr. Martha Escobar for their psychology expertise, and Faith Patterson, Debbie Price, Eric Coker, and Tim Murphy for their help in making this project possible.

REFERENCES