Towards Defining Dental Drilling Competence, Part 1: A Study of Bone Drilling Technique


Abstract: Technical skills are critical for dentists. Computer-based simulation offers a range of potential benefits for surgical training, but to date the development of simulators has not been characterized by a structured investigation of specific mechanisms by which trainees attain competence. This two-part study contributes to the understanding of the manner in which surgical psychomotor skills are acquired so that this knowledge can be incorporated into the design of training simulations. We studied participant groups of varying skill levels as they performed a drilling task in oral surgery. In this first part of our study, we investigated the elements of surgical technique and differences in the drilling performance of novice, competent, and expert dentists. Our results indicate that novice dentists employ a technique that differs considerably in drilling stroke length and duration from that employed by experts. Expert dentists perform faster, apply more force, lift the burr off the bone less, and produce superior results compared with novices.

Ms. Ioannou is a Ph.D. student, Department of Computer Science and Software Engineering, University of Melbourne; Dr. Kazmierczak is Senior Lecturer, Department of Computer Science and Software Engineering, University of Melbourne; Dr. Stern is Senior Lecturer, Department of Computer Science and Software Engineering, University of Melbourne; Dr. Smith is Head of the School of Dentistry, Director of the Oral Health Centre of Western Australia, and Winthrop Professor, University of Western Australia; Dr. Wise is Senior Lecturer, Social Science Group, Swinburne University of Technology, Victoria, Australia; and Dr. Field is Associate Professor, Department of Mechanical and Aerospace Engineering, Monash University, Victoria, Australia. Direct correspondence and requests for reprints to Ms. Ioanna Ioannou, Department of Computer Science and Software Engineering, University of Melbourne, 111 Barry Street, Carlton, Victoria 3053, Australia; 61-402-385-439 phone; 61-3-9348-1184 fax; ioannoui@unimelb.edu.au.

Keywords: dental education, preclinical dental training, oral surgery, psychomotor skills, technical expertise, virtual reality simulation

Submitted for publication 12/9/09; accepted 3/2/10

Dentists are required to possess excellent technical skills in order to perform surgical tasks quickly, accurately, and without undue trauma to the patient. However, dental curricula often provide limited hands-on technical skills training. As a consequence, the learning of practical skills occurs during clinical practice. This model of training does not guarantee adequate practice in key surgical skills, is not resource-efficient, and can put patients at risk. Some constraining factors leading to inadequate practical experience are lack of control over the variety of situations for practicing surgical tasks and limited access to skilled dentists who can train students.

A number of adjuncts to surgical skills training have been suggested to give students a better preparation for clinical practice. Among these is the use of virtual reality (VR) simulators to provide more practice opportunities and greater exposure to the broad range of conditions met in clinical practice. Existing studies show good evidence for skill transfer and good evidence for validity when VR simulators are used as adjuncts to traditional training.1

The goal of this study is to improve the design of VR simulations for dental training by investigating the skill acquisition process. Understanding how surgical motor skills are acquired will enable us to design simulators that have the ability to promote competence and are more closely aligned with the training goals at each stage in the training continuum. To this end, we have studied the differences between expert, competent, and novice dentists with the aim of characterizing those factors that distinguish expert behavior from novice behavior.
This study is presented in two parts. Part 1 (this article) investigates the factors that differentiate the technical ability and operative performance of predoctoral dental students and practitioners as they perform a basic oral surgical drilling task: removing the bone surrounding a tooth without damaging the tooth itself. We explore the implications of these differences for the design of training simulators for oral surgery. Part 2 of this study (reported in the accompanying article) examines the psychological aspects of competence by investigating the psychomotor cues used in decision making by various participant groups.

**Technical Skill Training in Oral Surgery**

In oral surgery, the training of technical skills begins with practice on synthetic replicas of human teeth or animal jaws, before moving on to supervised interactions with patients. A significant part of dental training follows the traditional apprenticeship model and consists of clinical work with patients. In the apprenticeship model of training, learners observe demonstrations of a task, then practice with coaching until they are able to perform independently. Deliberate effortful practice has been recognized as a crucial aspect of developing expertise. However, current dental training lacks sufficient opportunities to practice, as well as adequate feedback during and after practice.

**Dental Curricula**

Learning through practice is important in the acquisition of surgical motor skills. Durham et al. highlighted the lack of opportunity for practice in dental curricula in the United Kingdom when they found that dental surgery students at the University of Newcastle were exposed to a mean of only four tooth extractions over their entire bachelor’s degree. Henzi et al. studied the strengths and weaknesses of dental training in the United States from the student’s point of view. In their studies, some of the most commonly reported weaknesses of dental training included lack of exposure to the clinic early in the curriculum, ethically questionable treatment of patients, shortage of teaching staff, and lack of opportunity to learn new technologies and techniques. The studies by Henzi et al. highlight the difficulties involved in the existing clinic-based training methods and the need for improved surgical skills training.

Recently, there has been a rethinking of the methods used in surgical skills training and an interest in the augmentation of clinic-based training with extraoperative practice in surgical skill laboratories. In such laboratories, trainees use extraoperative training adjuncts to acquire and practice technical skills. Surgical skill laboratories in dentistry offer practice on animal jaws (porcine or ovine, since these offer a close resemblance to the density of the human jaw and teeth) or mannequins with synthetic teeth and jaws that are designed to closely resemble the look, structure, and feel of patients. Training on animal jaws and synthetic jaws both have their drawbacks, including high costs, difficulty in providing sufficient opportunities for practice, and low fidelity. These drawbacks make animal jaws and synthetic models a costly and inefficient option for regular practice.

**VR Simulation as a Training Adjunct**

VR simulation is a promising area for surgical training. A range of potential benefits have been proposed, including the opportunity for practice away from the stress of the operating room; the opportunity to make mistakes without harming patients; the ability to practice a variety of surgical tasks using the same equipment; immediate and objective performance feedback; and round-the-clock availability to students without need for constant attendance by faculty staff or replacement of physical materials.

A small number of virtual reality and augmented reality systems have been developed for dental training, including DentSim (Image Navigation Ltd., Jerusalem, Israel), the Virtual Reality Dental Training System (Novint Technologies, Albuquerque, NM), the Iowa Dental Surgical Simulator (College of Dentistry, University of Iowa, Iowa City, IA), PerioSim (College of Dentistry, University of Illinois at Chicago, Chicago, IL), and the VOXEL-MAN apicectomy simulator (University Medical Center Hamburg-Eppendorf, Hamburg, Germany). To our knowledge, only DentSim has been deployed in dental training, while all other systems are still being developed or tested. Evaluation of these systems has been encouraging. LeBlanc et al. and Buchanan evaluated the effectiveness of DentSim in training with positive results. Both studies found that the performance of students improved when using Dent-
Sim. Steinberg et al.\textsuperscript{2} found that experienced dental faculty members were highly enthusiastic about the potential of the PerioSim simulator for aiding the development of basic procedural skills. Sternberg et al.\textsuperscript{13} studied the transfer of skills learned using the VOXEL-MAN VR apicectomy simulator to porcine models and found that the group that received simulator training had higher probability of preserving vital neighboring structures and improved ability to objectively self-assess their performance compared to the group that did not receive simulator training. Whilst these studies are encouraging, more work is needed to test the transferability of skills learned through such simulators to the patient clinic.

**Limits of VR Simulation**

One of the recurrent problems with VR-based training systems is that their development has been largely technology-driven. Technology alone may or may not have instructional features because it is often employed without the benefit of findings from the science of training.\textsuperscript{16} Fiore et al. indicate that the progression from novice to expert has not been sufficiently examined.\textsuperscript{17} The abilities and needs of learners at each level of expertise must be understood, and training must be designed accordingly.

Few extraoperative training adjuncts in medicine have been developed with careful consideration for the training needs of their users and the requirements of the skill or task being trained. The exceptions are the Minimal Invasive Surgical Trainer (MIST) simulator for minimally invasive surgery\textsuperscript{11} and the Integrated Environment for the Rehearsal And Planning of Surgical Intervention (IERAPSI) temporal bone simulator. Both of these systems were developed after conducting task analyses that helped identify the training requirements that these systems had to address.\textsuperscript{18}

A similar analysis of training requirements would be valuable in developing training adjuncts for dentistry and oral surgery. Durham et al.\textsuperscript{8} emphasize the lack of studies of competence in oral surgery and the need to analyze students’ needs, identify their training requirements, and provide an environment conducive to achieving competence. Qualitative research in dental training such as that of Fugill\textsuperscript{7} only begins to address the gap, and the body of literature is still very limited. Clearly, more work is needed to understand the development of competence and the requirements that training tools must satisfy in order to facilitate skill acquisition.

This study addresses the need to understand competence and training simulator requirements in oral surgery in part by investigating the physical and cognitive aspects of competence with respect to a specific oral surgery skill. We compare study participants at three ability levels to identify differences in performance outcomes, differences in performance characteristics, and differences in the cues used for decision making during the chosen surgical task.

**Study Design**

This study is exploratory in nature. We employ a combination of quantitative and qualitative techniques to gain initial insights into the salient factors in surgical performance. In turn, these insights can inform the development of training strategies and the design of subsequent quantitative studies of such factors.

The study focuses on understanding the cues and factors involved in the acquisition of a particular practical skill in oral surgery and on characterizing expert and novice performance. We chose to study the skill of distinguishing bone from tooth during drilling because it is a general skill required in many oral surgery procedures.

To carry out our study, a suitable drilling task had to be chosen to demonstrate how well participants can distinguish tooth from bone. The selected drilling task was that of removing bone to expose the root of a tooth without damaging the tooth itself. The task was carried out on an ovine jaw using a Surgairtome Two drill (Hall Surgical, Largo, FL) with a fissure bur. This task was chosen because it requires the ability to differentiate tooth and bone in order to remove only the appropriate amount of bone and avoid damaging the tooth, such that extraction can be carried out cleanly and adjacent teeth are not damaged. It is a simple task to teach to all participant groups, but still sufficiently complex to require specialized psychomotor skills. Furthermore, the result of the task can be rapidly assessed.

The participants were dental students at the University of Melbourne and practicing dentists. Participants were separated into groups based on Hoffman et al.’s expertise classifications.\textsuperscript{19} Table 1 summarizes the participant numbers and characteristics of each group, designated as initiates, journeymen, and experts. Ethics approval from the home institution of the lead author was acquired for the use of human subjects in this study.

---

*September 2010 • Journal of Dental Education*
Data Collection Methods

The experimental procedure was the same for all participant groups. Participants watched an instructional video explaining the task and its goals and providing a demonstration of good and bad performance. This was followed by a brief interview in which they were asked to explain how they planned to perform the task and the cues that they were going to look for. They then performed the task. Upon its completion, they responded to another interview examining the cues and factors that affected their actual performance in the task and completed a written questionnaire. Part 1 of the study focuses on the physical aspects of performance, while Part 2 presents the results of the analysis of cues and factors based on interview and questionnaire results.

Audio and video of the session were recorded using two video cameras at different positions. One camera recorded the entire scene during the session, while the second camera recorded a close-up of the jaw during the task. The forces applied during drilling were recorded using a custom-built tri-axial force sensor composed of three iLoad Digital USB load cells (Loadstar Sensors Inc., Fremont, Canada) to which the jaw piece was attached. Figure 1 illustrates this setup.

Data Analysis Methods

As a first step in identifying differences between experts and initiates, we analyzed the characteristics of their drilling technique. This was done by qualitatively coding the close-up videos of the drilling task for each participant to obtain a number of measurements, as outlined in the following sections. No predefined coding scheme was used. Instead, the coding scheme was developed based on initial observations of the data and analysis of the patterns that emerged with the guidance of one of the coauthors (ACS), who has twenty-three years of experience in the area.

Drilling stroke analysis was an important part of data analysis. A stroke was defined as a motion of the bur across the surface of the bone with no significant change in direction. Strokes were classified as either leftwards or rightwards, based on the overall direction of the bur motion across the surface of the bone in relation to the participant.

Force sensor readings were quantitatively analyzed to derive the average and maximum force magnitudes applied by different groups. The direction of the forces applied was also analyzed and compared to the results of our qualitative stroke analysis.

Finally, the drilled jaw specimens were graded by an expert assessor blind to the participant’s identity and level of prior experience. The assessor used a scale of 1 to 10, with 10 being the best outcome. The grading criteria used were 1) appropriate amount of bone removal to clearly expose the surface of the root and 2) minimal amount of tooth drilled.

Table 1. Participant groups, numbers of participants, and group characteristics for this study

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Conceptual Definition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiate</td>
<td>5</td>
<td>A novice who has begun introductory instruction.</td>
<td>Dental science students in their fourth year of study. By this stage they have performed under five similar tasks in a skill laboratory and have been training in the clinic with patients for approximately six months.</td>
</tr>
<tr>
<td>Journeyman</td>
<td>5</td>
<td>An experienced and reliable worker, or one who has achieved a level of competence.</td>
<td>Practicing dentists with six months to ten years of dental experience, who were undertaking postgraduate study in oral surgery.</td>
</tr>
<tr>
<td>Expert</td>
<td>4</td>
<td>A distinguished journeyman, highly regarded by his or her peers, whose judgments are uncommonly accurate and reliable, whose performance shows consummate skill and economy of effort, and who can deal effectively with rare or “tough” cases.</td>
<td>Practicing dentists and oral surgeons with more than ten years of experience in oral surgery. Ten years is the period of practice that is typically required to attain expertise in an area.</td>
</tr>
</tbody>
</table>

Sources:
Results

We begin by comparing the training task outcomes of each group to examine overall performance. We then explore differences in performance by comparing drilling technique and other performance characteristics across groups.

Task Outcomes

Table 2 summarizes the scores assigned to participants by the expert assessor based on the end result of the task. High variance was noted within each participant group. Journeymen and experts had a higher minimum score and higher median score than initiates. Experts had significantly higher average scores than initiates ($t=1.77$, df=7, $p=0.06$), while the difference in the average scores of initiates-journeymen and journeymen-experts was not statistically significant.

The time taken to complete a task is sometimes used as an indicator of performance. The time spent in carrying out the task was measured from the moment the bur touched the bone to the moment the bur was lifted off the bone upon completion of the task. This time includes pauses made during the task. Table 3 summarizes the results for each group. The time it took to perform the task steadily decreased across the initiates, journeymen, and experts, with the last two groups being quite close in average time.

---

**Table 2. Task outcome scores out of a maximum of 10 based on drilled jaw specimen**

<table>
<thead>
<tr>
<th>Group</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiates</td>
<td>0</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Journeymen</td>
<td>3</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Experts</td>
<td>5</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

---

**Table 3. Time taken to perform task across groups (in seconds)**

<table>
<thead>
<tr>
<th>Group</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiates</td>
<td>58</td>
<td>72</td>
<td>62</td>
</tr>
<tr>
<td>Journeymen</td>
<td>22</td>
<td>44</td>
<td>28</td>
</tr>
<tr>
<td>Experts</td>
<td>12</td>
<td>30</td>
<td>25</td>
</tr>
</tbody>
</table>
taken. On average, initiates took significantly more time to complete the task compared to journeymen ($t=7.40, df=8, p<0.001$) and experts ($t=9.12, df=7, p<0.001$).

The remaining sections present our analysis of the physical characteristics of the drilling technique employed by participants.

**Stroke Strength and Direction**

The forces applied to the jaw during the task were measured in three directions: X, Y, and Z as shown in Figure 2. Force data from each of the three force sensors were recorded every 0.015 seconds during the task, resulting in a set of 3D force points for each participant, indicating the amount of force applied in each direction X, Y, and Z at each point in time. Each force point is comprised of three directional forces. The total force magnitude is calculated using Equation 1. The average force magnitudes and ranges for each participant group were calculated based on the set of force points collected for each participant. The force sensor had a uniform error of ±0.2N. Due to inherent uncertainty in the calibration...

\[ \| F \| = \sqrt{\langle F_x \rangle^2 + \langle F_y \rangle^2 + \langle F_z \rangle^2} \]

**Equation 1: calculation of force magnitude for each force point**

<table>
<thead>
<tr>
<th></th>
<th>$| F_x | \quad \langle \text{Range} \rangle$</th>
<th>$| F_y | \quad \langle \text{Range} \rangle$</th>
<th>$| F_z | \quad \langle \text{Range} \rangle$</th>
<th>Total $| F | \quad \langle \text{Range} \rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiates</td>
<td>0.2 (max 0.6)</td>
<td>0.1 (max 0.5)</td>
<td>0.2 (max 0.6)</td>
<td>0.3 (max 1.0)</td>
</tr>
<tr>
<td>Journeymen</td>
<td>0.1 (max 0.4)</td>
<td>0.1 (max 0.5)</td>
<td>0.1 (max 0.4)</td>
<td>0.2 (max 0.8)</td>
</tr>
<tr>
<td>Experts</td>
<td>0.2 (max 0.8)</td>
<td>0.3 (max 1.0)</td>
<td>0.2 (max 0.7)</td>
<td>0.4 (max 1.5)</td>
</tr>
</tbody>
</table>

**Figure 2. Force directions with respect to participants**

*Note:* From the participant’s point of view, X is backwards and forwards, Y is leftwards and rightwards, and Z is upwards and downwards.
tion of the sensor, the force measurements may not be exact on an absolute scale. However, they can be used to make relative comparisons, as was required for this study.

Table 4 shows the average forces applied by each group. Overall, journeymen applied the least force compared to experts (t=3.34, df=7, p=0.006) and initiates (t=2.03, df=8, p=0.04). Experts applied slightly more force than initiates on average (t=1.68, df=7, p=0.07). The total average force was 0.2–0.4N with average maximum of 0.8–1.5N depending on group. This data indicate the approximate range of forces applied for a task of this type.

Whilst the magnitude of the forces is similar across groups, the direction of the forces applied varies widely. Figures 3, 4, and 5 show three-dimensional (3D) plots of the force points for each participant in each group. Different color shades in each graph designate different participants. As Figure 3 shows, force points from initiates were mostly in the negative Y direction (i.e., leftward), negative X direction (i.e., inward), and positive Z direction (i.e., downward). Journeymen (Figure 4) applied the most force along the Y direction (both leftwards and rightwards) but overall they applied very little force; thus, their force points concentrate around the origin. By contrast, expert force points (Figure 5) are more widely spread around all axes. Individual experts appear to have different tendencies, and the overall pattern is that they apply more force than any of the other groups.

A qualitative stroke analysis was performed to identify drilling stroke characteristics. The total number of leftwards and rightwards strokes was counted for each participant, thus enabling us to calculate the ratio of rightwards to leftwards strokes. Initiates were found to have a bias towards rightwards strokes compared to journeymen (t=1.42, df=8, p=0.10) and experts (t=1.47, df=7, p=0.09). A bias towards
rightwards strokes would result in a bias towards negative Y forces, as reflected in Figure 3. Thus, the stroke analysis data support the force sensor findings. Journeymen had a ratio of 1:1, and experts had a ratio of 0.8:1, indicating that both these groups are much closer to a balanced stroke technique, which would be represented by a ratio of 1:1.

**Stroke Duration**

The average stroke duration for each participant was calculated by counting the total number of strokes and averaging it across the time it took to perform the task minus the duration of any pauses (where a pause is defined as a lift of the bur that lasts at least 1.5 seconds). Figure 6 shows the average time per stroke for each participant group. It is evident that the expert participants used distinctly longer strokes than initiates ($t=4.55$, df=7, $p=0.002$) and journeymen ($t=8.60$, df=8, $p<0.001$).

Close-up video observation of expert participants showed long sweeping strokes from one side of the tooth to the other, which is consistent with Figure 6. Initiate participants used primarily short jabbing strokes whereby they lifted and repositioned the drill after each stroke in a circular motion. Figure 7 demonstrates the two types of strokes observed. The jabbing technique of initiates may be responsible for the rightward stroke bias discussed previously.

Journeymen generally performed somewhere in between, with some long sweeping strokes and some jabbing strokes. However, unlike initiates, journeymen did not frequently lift the bur between strokes and did not use circular motion. It should be noted that the initiate represented in Figure 6 by the point with average stroke time of 0.825 did not actually use long strokes, but rather lifted the drill for longer periods between strokes. This value is an artifact of the stroke analysis method rather than a valid result.

**Bur Lifts and Pauses**

As mentioned above, initiates lifted the bur off the bone very frequently. Bur lifts were counted by identifying all individual instances in the video in which the bur was lifted off the bone for less than 1.5 seconds. Bur lifts of more than 1.5 seconds were classified as pauses. Figure 8 shows the number of bur lifts for each participant group. We observed that the majority of journeymen and experts lifted the bur far less than initiates ($t=3.35$, df=8, $p=0.005$ and $t=2.74$, df=8, $p=0.014$ respectively). The fact that journeymen and experts lifted the bur far less than initiates ($t=3.35$, df=8, $p=0.005$ and $t=2.74$, df=8, $p=0.014$ respectively) supports the idea that they use a more balanced stroke technique.
men and experts keep the bur on the bone as much as possible suggests more efficient drilling, since the bur is removing material at all times.

Finally, all but one initiate paused during the task with an average of 2.2 pauses. Only a single journeyman paused once, while all experts completed the task without any pauses.

Discussion

VR training tools with force feedback are expected to provide more opportunities to practice than existing training methods. Our technique analysis has identified the approximate forces applied in removing bone around teeth. These forces range between 0.2 and 0.4N on average with maximums of 0.8 to 1.5N (Table 4). These force estimates can be used to calibrate the force feedback of a visuo-haptic simulation, such that the forces experienced during the simulation are within a realistic range.

The technique analysis carried out has revealed important differences between experts and novices. Stroke analysis indicated that experts use longer sweeping strokes when drilling (Figure 6), do not frequently lift the bur off the bone (Figure 8), and do not pause. By contrast, initiates used short jabbing strokes with a bias in one direction, frequently lifted the bur, and frequently paused. Initiates paused during the task, while only one journeyman and no experts paused. One possible explanation for this difference is that initiates may have felt the need to pause and visually review their progress, whereas journeymen and experts had a better awareness of the effects of their past bur strokes and thus no need to pause. The experts' technique is more efficient and thus more desirable because the drill is constantly removing material. Experts also applied the largest forces compared to the other groups, which may indicate more confidence.

Understanding these differences is important because it enables the identification of characteristics of good technique and the development of metrics to measure performance. Such metrics can be incorporated into a simulation tool to enable automated performance feedback and objective assessment. For example, a simulator can determine the degree to which a trainee is using short jabbing strokes instead of more efficient sweeping strokes and provide feedback as the participant is carrying out the task. Force measurements are another possible metric, which can be used to alert a trainee when too much or too little force is applied. Moreover, a simulator can provide objective assessment of the task outcome, by calculating the exact amount of bone material removed and determining the extent to which tooth was removed.

This study has provided an initial understanding of the characteristics of expert performance in oral surgery and some shortcomings in student performance, which may be utilized to inform the development of simulators for technical skills training in oral surgery. The factors identified in this study should be further examined to determine the trend in developing expert drilling skills—that is, whether metrics such as stroke length and the number of bur lifts are true predictors of expert performance. What is important here is to establish if there is a modal change involved in expertise or whether expertise is developed incrementally. The current study suggests areas of possible significance in the development of expertise in dental drilling, but a larger study is essential in improving our understanding. Also, a larger study of the forces applied during various drilling tasks is required to create a basis for comparison between the forces experienced in simulators and the forces experienced in real drilling.

Conclusion

This study represents a novel attempt to analyze the elements of surgical technique and characterize expertise in the domains of dentistry and oral surgery. With improved understanding of the characteristics of expertise, dental educators can better advise students on how to improve their performance, design better tools for practical skills development, and provide objective assessment. Studies of this type have been carried out in other fields such as sports and aviation and other areas of medicine. However, to date, dentistry and oral surgery lag behind. Studies of this type in the domain of oral surgery can inform the design of effective and efficient tools to train expert dentists and oral surgeons.

This first part of our study has focussed on the physical aspects of technique across different skill groups. Part 2 will focus on the cognitive aspect of performance and the psychomotor cues used by participants for decision making during a drilling task.

REFERENCES


