Palpatory Diagnosis Training on the Virtual Haptic Back: Performance Improvement and User Evaluations

John N. Howell, PhD; Robert R. Conatser, MS; Robert L. Williams II, PhD; Janet M. Burns, DO; and David C. Eland, DO

Context: Learning palpatory diagnosis is a challenge for many osteopathic medical students. The Virtual Haptic Back (VHB) is an aid in teaching and learning these skills. The device simulates the contours and surface compliances of the human back and allows these to be felt through haptic interfaces. Regions of abnormal tissue texture are simulated by altered surface compliance.

Objectives: To examine the effectiveness of the VHB in training osteopathic medical students in palpatory diagnosis and to determine students’ subjective impressions of the potential value of the VHB in learning palpatory diagnosis.

Methods: Twenty-one first-year osteopathic medical students at the Ohio University College of Osteopathic Medicine in Athens took performance tests on the VHB before and after a series of eight 15-minute practice sessions, which occurred during multiple 2-week sessions between September 2005 and January 2006. The tests and practice sessions measured student accuracy and speed in locating regions of abnormal compliance in the haptic back, which varied randomly among 12 sites in the thoracic region. After completing the practice sessions and performance tests, students filled out a questionnaire regarding their impressions of the potential value of the VHB as a learning aid.

Results: Students collectively improved in accuracy and speed following the practice sessions compared with the initial performance test. Subjects improved from being able to detect only a 40% difference in compliance to being able to detect an 11% difference (P<.05). The greatest improvements occurred at the difficulty levels near the apparent detection limit of compliance differences. Survey responses indicated that students thought the VHB experience was helpful in developing their palpatory skills.

Conclusion: The VHB has potential as an effective aid to osteopathic medical students in learning palpatory diagnosis.

Palpation of the human back is used diagnostically by osteopathic physicians and other clinicians to detect musculoskeletal abnormalities collectively referred to as somatic dysfunction. One characteristic of these abnormalities is altered tissue texture, which reflects altered tension in underlying muscles and connective tissues. A key component of tissue texture is tissue compliance (tissue displacement per unit of applied force). The reciprocal of compliance is elastance, more commonly known as stiffness (force generated per unit of displacement).

The Virtual Haptic Back (VHB) is being developed at the Ohio University College of Osteopathic Medicine (OU-COM) in Athens as an aid to the teaching and learning of palpatory diagnosis. The term haptics refers to the human sense of touch. The VHB simulates the contour and compliance properties of the human back. A detailed description of this new technology is included in the Appendix.

The smallest difference that can be detected by any sensory system of the body is called the “just-noticeable difference.” When this difference is expressed as a fractional change, it is known as the Weber fraction. The Weber fraction is sometimes expressed as a percent. For example, a Weber fraction of 0.11 indicates that one can detect an 11% difference. In a study of compliance detection, Dhruv and Tendick, using a PHANTOM 1.5 haptic interface (SensAble Technologies Inc, Woburn, Mass) found Weber fractions in the range of 0.14 to 0.25 for a simple mechanical task consisting of pressing a finger against a resistance that behaves as a linear spring.

In a similar study, DeGersem reported Weber fractions between 0.08 and 0.12 achieved by 6 subjects. The range of compliance values used by DeGersem included the range of compliance values previously measured over paraspinal muscles (0.8 mm/N to 1.2 mm/N in the thoracic region and up to 1.6 mm/N in the lumbar region).

The VHB, however, presents students with a more complex task than those studied by Dhruv and Tendick and...
DeGersem. With the VHB, the location of the abnormalities on the back is varied randomly and the difficulty level of the task can be varied by making the abnormalities obvious or very subtle. In the current study we sought to determine if training on the VHB would increase the ability of users to detect small differences in compliance between adjacent areas on the back. The current study shows that, through practice, VHB users were able to achieve the same level of compliance discrimination that had been achieved by subjects in the previously described studies.3,4

Methods
Subjects (N=21) were first-year osteopathic medical student volunteers from OU-COM. All subjects were within the first 5 months of their palpatory training. The proposal for this research was submitted to the Ohio University Institutional Review Board and was judged to be exempt from review.

Preliminary Session and Pretest
During their first session in the laboratory, subjects were given an opportunity to familiarize themselves with the haptic interfaces, practicing 10 to 15 minutes to identify regions of abnormal compliance. During this familiarization period, a transparency function was activated (Figure 1), permitting the subject to see the skeletal elements beneath the skin for reference.

Following this preliminary phase, subjects took a test with an opaque virtual back (Supplemental Figure 2), in which they had to locate the regions of abnormal compliance presented in successive trials. The location of the abnormal region of the haptic back varied randomly among trials. The abnormalities could be on either the left or right side and at any one of six thoracic levels, T5 through T10 (Figure 1). Subjects typically moved their fingers along the haptic back searching for regional differences in tissue compliance, later returning to regions they suspected might be abnormal. When they had decided which area was abnormal, they pressed a foot switch while holding a finger on the abnormal area. The system provided immediate verbal feedback as to whether their choice was correct or not.

Five levels of difficulty (0.5, 0.7, 0.8, 0.9, and 0.95) corresponding to various compliance differences and Weber fractions were presented in the pretest (Table). The pretest began with the easiest level (0.5) and progressed incrementally to the most difficult level (0.95). Two trials occurred at each difficulty level. Subjects were required to complete each trial within 1 minute (time remaining was visible on the screen). Midway through the test, the program paused, giving the subject an opportunity to take his or her fingers out of the apparatus and rest.

Practice Sessions
Immediately following the pretest, subjects completed the first of eight practice sessions to be finished in 2 weeks. Subjects were permitted to complete the practice sessions at their own convenience but were allowed no more than one session per day. Individual trials during the sessions were not limited or timed, but each practice session was limited to 15 minutes.

Nine levels of difficulty were available in each practice session compared with the five levels used in testing (Table). Although the default setting was the easiest level (greatest compliance difference) for each practice session, subjects could pick any level of difficulty at which to work. Most started with the easier levels in the earlier sessions and progressed to the harder levels in the later sessions.

In the practice sessions, when subjects incorrectly identified an abnormality in tissue compliance, the program alerted them of their error and displayed a box around the correct

| Table Compliance Values and Corresponding Weber Fractions for Each Difficulty Level in the Virtual Haptic Back* |
|-----------------|-----------------|-----------------|
| Difficulty Level | Compliance, mm/N | Weber Fraction, % |
| 0               | 0.97            | 61.5            |
| 0.25            | 1.19            | 52.8            |
| 0.5†            | 1.51            | 40.1            |
| 0.7†            | 1.98            | 21.4            |
| 0.75            | 2.14            | 15.1            |
| 0.8†            | 2.25            | 10.7            |
| 0.85            | 2.29            | 9.1             |
| 0.9†            | 2.35            | 6.7             |
| 0.95†           | 2.45            | 2.8             |

* Background compliance was 2.52 mm/N.
† Difficulty level used in the pre- and posttests.
area on the screen with the transparency function turned on. Subjects could then feel the abnormality before going on to the next trial. Subjects could also pause the practice session at any time to rest.

**Posttest and Survey**

To evaluate students for performance improvements, the pretest previously described was re-administered as a posttest at the end of the 2-week practice sequence. After completion of the posttest, students were asked to complete a brief survey regarding the VHB and its value as a learning tool.

**Data Analysis**

Results from performance tests and practice sessions were analyzed with a repeated-measures analysis of variance (ANOVA). A Bonferroni posthoc analysis was also performed on the results of the practice sessions. Survey responses were recorded and totaled.

**Results**

**Performance Tests**

Student accuracy of localizing the dysfunctional areas was statistically significant \( P < .05 \) at difficulty levels 0.7, 0.8, and 0.9 between the pretest, taken before the eight practice sessions, and the posttest, taken after the practice sessions (Figure 2). No statistically significant improvements were seen at the easiest level, presumably because performance at this level was quite high even in the pretest. Likewise, no statistically significant improvements were seen at the most difficult level that was tested (0.95), presumably because the task was simply too hard for students to master. At this level, performance remained at near-chance levels. Speed improved at all levels of difficulty (Figure 3).

**Practice Sessions**

Analysis of the data from the practice sessions posed problems resulting from the fact that not all subjects chose to practice the sessions in order of easiest to most difficult. After the first few sessions, some subjects skipped the easier levels and went directly to the harder levels. Other subjects stuck primarily with the less difficult levels at which they felt competent. This resulted in missing data points both at the easier levels and at the more difficult levels.

We approached this absence of data in two different ways. First, we filled in missing data points. For example, if subjects had shown they were able to locate abnormalities in the easy levels correctly in early practice sessions, they were credited with doing them correctly in later practice sessions—even if they did not actually do them. At the more challenging levels, where subjects had not made correct localizations, incorrect responses were calculated even if they had not tried those levels. Second, we eliminated subjects with missing data points, which brought the total number of subjects down from 21 to 14. Analysis of the data is presented both ways in Figure 4 and Figure 5, respectively.
Data from the practice sessions for all 21 subjects are shown in Figure 4. At the easiest levels, users performed very well even in the first practice session. At the most challenging level, student accuracy improved with successive practice sessions, but remained at near-chance levels throughout. The most dramatic improvement is seen at the intermediate levels of difficulty. Subjects performed poorly during the first practice sessions, but did progressively better in successive sessions. Repeated-measures ANOVA indicated statistically significant differences (P<.05) across visits and across difficulty levels as well as a significant interaction term. Bonferroni posthoc analysis indicated the following:

- The results of the first practice session are statistically different from the results of all subsequent sessions.
- The second and third practice sessions are statistically different from later sessions.
- The fourth through eighth sessions are not statistically different from each other.

Although these observations generally corroborate the impression one gains from Figure 4, the requirements of repeated-measures ANOVA with respect to sphericity and normality of the data set were not met.

As previously mentioned, and to meet ANOVA requirements, we analyzed a subset of the data using results from the 14 subjects who tried all nine difficulty levels, starting with the easiest up to the hardest levels they attempted in each session, leaving no missing data points (Figure 5). The same general pattern is evident in this subset as was seen with the entire population (Figure 4), and repeated-measures ANOVA revealed statistically significant differences between performance during the first four sessions compared with the last four (P<.001).

There was also a statistically significant interaction period (P=.024), indicating that the change in performance between the first and last group of sessions varied with difficulty level (ie, only at the harder levels did improvement occur with practice).

Survey

Students responded to a brief survey after completing the VHB posttest. Of the 21 subjects, 17 (81%) marked “yes” and 4 (19%) marked “maybe” in response to the question, “Do you think this practice with the haptic back will be of help to you in the development of your palpatory skills in OMM lab?” None of the students answered “no.”

Twelve students (57%) agreed that further practice with the haptic back would be of help in the development of their palpatory skills, 8 (38%) answered “maybe,” and 1 (5%) responded “no.”

Using a Likert scale, the students were also asked to rate the realism of the simulation, with 0 being “unrealistic” and 10 being “realistic.” The mean rating was 6.5.

Comment

Both the objective results obtained in this study and the subjective responses of students indicate the potential value of the VHB as an aid in learning palpatory diagnosis. Data from the pre- and posttests indicate statistically significant skill improvement (P<.05), and the data from the practice sessions reveal the pattern of improvement. The students in the study were first-year osteopathic medical students taking a course in OMM at OU-COM, where OMM training occurs over the course of 2 years and consists of 2 hours of practical training weekly in a laboratory supplemented with occasional lectures. The VHB study was carried out during the fall and winter quarters in 2005 during the early stage of students’ palpatory training.

Some of the subjects’ skill improvement between the pre- and posttest undoubtedly resulted from familiarization with
been done on the ability of subjects to detect differences in compliance of real objects\(^7\) and of virtual objects.\(^3,4\) As previously discussed, compliance detection of virtual objects or surfaces has been done with the PHANTOM haptic interface and has yielded Weber fraction estimates as low as 0.08 to 0.12 with time-invariant surfaces.\(^4\)

In the current study, the smallest compliance difference detectable is judged by correct localization of the abnormal area of the back, so the chance value is far less than 50%. The actual area that is abnormal, 7.5 cm\(^2\), constitutes only 2.5% of the area in which the palpation is done. However, there are only 12 different sites where the abnormality can occur in the VHB. Assuming the students knew the location of those twelve sites and were applying only one finger for identification of the abnormal area, the chance level would be 1 in 12, or 8.3%. However, with two-finger palpation, the student could have been touching two different areas at once when he or she pressed the foot switch. If either finger is on the correct area, the user is credited with a correct answer. That could, in principle, raise the chance level to 1 in 6, or 17%. Thus, chance level may have varied among students depending on their approach. We have chosen 20% as a conservative estimate of chance level, though it must certainly be lower than that.

Using 20% as chance level, we can take the performance level of 60% as the just-noticeable difference (again, halfway between chance [20%] and certainty [100%]). During the pretest, the 60% criterion for the just-noticeable difference was achieved only at the easiest tested level, difficulty level 0.5 (imagine a horizontal line at 60% representing master level in Figure 2). This corresponds to a Weber fraction of 0.40. During the posttest, the criterion was met at the 0.8 level, corresponding to a Weber fraction of 0.11. During the first practice session, performance at difficulty levels of 0.7 or less reached the criterion level, corresponding to a Weber fraction of 0.21. During the sixth practice session, the 60% level was achieved at the 0.85 difficulty level, corresponding to a Weber fraction of 0.09. This fraction falls in the range obtained by DeGersem,\(^4\) who used a standard psychophysical design in which subjects palpated two smooth.

In typical psychophysical experiments designed to determine the sensitivity of any sensory system, a large number of trials is typically carried out, simply asking subjects to determine which of two inputs is larger than the other.\(^2\) In the case of detecting compliance differences, for instance, subjects are typically asked to compare the compliances of several different test objects with that of a reference object. When the compliances are very close to that of the reference object, subjects will not be able to distinguish between them and are correct only 50% of the time (chance level). When the differences are large, subjects are correct 100% of the time. Generally, the just-noticeable difference is taken as the level at which correct identifications are made 75% of the time (ie, halfway between chance and certainty). Experiments of this type have

Figure 5. Subject improvement in performance accuracy in the Virtual Haptic Back during the eight practice sessions and across nine levels of difficulty (n=14). *Difficulty level used in the pre- and posttests.
surfaces and determined which surface had the higher compliance. However, the task in our experiment was more complex in that (1) the areas of abnormal compliance first had to be searched for and found, and (2) the abnormal compliance was superimposed not on a flat surface, but on a surface with the complex contour of the human back. Further studies are needed to determine if a training effect, as we observed, would also be observed in the simpler experimental paradigm.

It is interesting that the data suggest performance improvement with successive practice sessions even at the 0.9 and 0.95 difficulty levels, with Weber fractions of 0.067 and 0.028. At these levels, the 60% mastery criterion was not reached, but the improvement seen raises the question as to whether further practice would have permitted students to reach that criterion level.

Several of the subjects were enthusiastic about using the VHB, indicating that it would have been particularly helpful even earlier in their OMM experience and recommending that all first-year students be given an opportunity to use it. Subsequently, in fall 2006, the VHB was incorporated into the OU-COM curriculum. Data collected from studying student use of the VHB in 2006 have been submitted for publication. Based on those data, a modified VHB was used as a required element in the curriculum in the fall of 2007.

It is our experience that osteopathic medical students often lament the paucity of supervised practice time in palpatory diagnosis and manipulative medicine, time that would allow for feedback about the correctness of their palpatory impressions and their application of osteopathic manipulative techniques. The immediate feedback provided by the VHB fills this need with regard to palpation, allowing students to develop confidence in their palpatory abilities. It also allows them to explore different modes of palpation (eg, use of different fingers to find out which work best for them).

The VHB represents an accurate model of palpation of the human back in the sense that the compliance values used in the model are in the range of those measured on human subjects and the palpatory forces are in the same range as those used by osteopathic physicians examining real patients. In its present state, however, the VHB simulates only the most rudimentary element of tissue texture change, namely a decreased compliance of tissues in a single area.

Patterns in real backs are far more complex and skill in palpatory diagnosis involves not only detection of tissue texture changes, but also interpretation of these changes in the context of complex patterns. Work is currently underway to make the VHB more realistic. One step is the programming of several haptic backs, reflecting palpatory differences with age, sex, and body habitus. Another step toward this goal is to incorporate some of the more complex patterns of tissue texture change, reflecting such elements as mirror image asymmetries and multiple segment interactions as described by William L. Johnston, DO. Work is also underway to permit the input of gross motion with one hand moving a simulated arm or shoulder, while the other hand palpates corresponding tissue texture responses in the back. These efforts are intended to extend the usefulness of the VHB simulation.

Conclusion

Eight 15-minute training sessions on the VHB permitted osteopathic medical students to improve their ability to discriminate compliance differences. The training effect, represented by the performance improvements in both speed and accuracy, coupled with the positive endorsements by student users, suggests that the VHB can serve as an effective teaching aid for palpatory diagnosis.

References

Approximately 15 cm behind the haptic back is a full-sized visual image of the back displayed on a 23-inch flat screen monitor (Supplemental Figure 2). Two dots on the screen, labeled “L” and “R,” visually indicate where the user’s fingers are with respect to the haptic back. In this way, the user is able to bring his or her fingers directly to the center of the haptic back in order to begin palpation. The VHB also has a transparency function, which reveals the skeletal elements of the back (Supplemental Figure 1), rather than just the standard opaque image (Supplemental Figure 2).

In the VHB model used in the current study, the back was programmed in C++ using the OpenHaptics software toolkit, GHOST SDK (SensAble Technologies Inc, Woburn, Mass), and OpenGL (version 1.5.1; SGI, Sunnyvale, Calif) for graphics. It was programmed to have a uniform compliance except for a small 2.5 cm by 3.0 cm area of simulated somatic dysfunction. The entire region of testing was a rectangle 13.5 cm wide and 22 cm high. The test region was superimposed on the graphics image of the back and encompassed thoracic segments T5 through T10.

(continued)
Appendix (continued)

The compliance of the abnormal region, which ranged from 0.97 mm/N to 2.45 mm/N in the current study, was made to blend smoothly into the compliance of the surrounding areas (2.52 mm/N) with the following hyperbolic tangent function:

\[ f(x) = \frac{1}{2} \left( \tanh\left( a(x-b)+c \right) - \tanh\left( a(x-b)-c \right) \right) \]

where \( a \) is the distance over which compliance transition occurs; \( b \) is the distance between the center of the abnormal area and a reference point, such as the body midline; and \( c \) is the width of the abnormal area. This adjustment prevents a sharp demarcation separating the abnormal regions from the normal regions.

In discriminating between two different linear compliances in the VHB, applying greater force causes increasingly greater differences in displacement. This difference initially leads users to press harder if they are having difficulty detecting the abnormal region. However, sustained application of force levels over 6 N can cause the electric motors of the haptic interfaces to overheat, which in turn can cause the program to shut down. The application of high forces is also clinically inappropriate because of potential patient discomfort and because palpatory information from superficial soft tissues can be lost. Higher force levels are used in palpation of the position of bony landmarks, such as transverse process of vertebrae. To discourage users from pushing too hard, we added the following components to the VHB:

- When users apply unacceptably high forces, automated voice feedback warns them not to press so hard.
- A visual gauge in the lower right corner of the screen monitors user force levels, enabling users to see when they are approaching unacceptably high force levels.
- Most importantly, the programmed compliance difference between the abnormal area and the surrounding areas is multiplied by a second hyperbolic tangent function, which makes the difference gradually disappear with increasing displacements between 8 mm and 16 mm.

Thus the compliance differences between the normal and abnormal regions were greatest, and therefore easiest to detect, when the user was applying palpatory force in the range of about 2.5 N to 3.5 N (Supplemental Figure 3). Based on preliminary measurements, this force level falls within the range of forces typically exerted by experts in clinical palpation diagnosis during palpation of superficial soft tissues.

References


Supplemental Figure 3. Relationship between the subjects’ applied force and displacement at different difficulty levels (N=21). The straight line indicates background elastance (compliance⁻¹). Increasing deviations from this background elastance make the task progressively easier. The deviations disappear at high displacements produced by application of high forces.