balt3/sih-sih/sih-sih/sih00410/sih0250-10z | xppws | S=1 | 6/18/10 | 8:08 | 4/Color Figure(s): F1,F3–11 | Art: SIH200270 | Input-us

AQ: A

A Stiffness Discrimination Experiment Including Analysis of Palpation Forces and Velocities

Ernur Karadogan, MS;
AQ: B
Robert L. Williams II, PhD;
John N. Howell, PhD;
Robert R. Conatser, Jr., MS

Introduction: The incorporation of haptics, the sense of touch, into medical simulations increases their capabilities by enabling the users to "feel" the virtual environment. We are involved with haptics-augmented VR training for palpatory diagnosis. We have developed a stiffness discrimination program to train and test users in finding subtle differences in human tissue stiffness for medical diagnoses. In this article, we studied the effect of surface stiffness on the stiffness discrimination task and analyzed the palpation force and speed during haptic exploration.
Methods: The ability to discriminate stiffness differences was studied by means of a psychophysical experiment with 13 second-year medical students (eight women and five men). Subjects were asked to identify the stiffer of two virtual computer-generated surfaces (top surfaces of two cylinders) by means of a PHANToM Omni (SensAble Inc.) haptic device with a modified stylus to accommodate their fingers. The modification of the stylus provided the mechanical advantage to simulate surface stiffness values that are beyond the original capability of the haptic device. An adaptive two-alternative forced-choice procedure was used on each trial. Palpation velocity and force vectors were recorded directly from the haptic device for further analyses. Weber fraction was determined by using an automated mastery algorithm.
Results: Four standard stiffness values (0.25, 0.50, 1.00, and 1.25 N/mm), typical of the stiffness range of human soft tissues, were used as references. The average experimental Weber fractions observed were 0.20, 0.27, 0.26, and 0.30, respectively, with higher Weber fractions corresponding to lower stiffness discrimination ability. At 1.00 and 1.25 N/mm standard stiffness, the correlation analysis for Weber fraction and the palpation speed revealed significant differences ($P < 0.05$). These differences suggested that the subjects with a higher palpation velocity tended to have a higher Weber fraction. There was no significant difference between male and female subjects. There was no significant difference between subjects new to the haptic device and those who had used it previously. The average amount of force that was applied by the subjects to the standard stiffness side and the comparison stiffness side within the sessions was not significantly different. However, the subjects increased the average force they applied with increasing standard stiffness value across the sessions ($P < 0.05$).
Conclusions: For the four standard stiffness values investigated, 0.25, 0.50, 1.00, and 1.25 N/mm, the resulting average stiffness-discrimination Weber fractions were 0.20, 0.27, 0.26, and 0.30, respectively. The average of the forces applied by the subjects was constant within a single session (with a single standard stiffness value). This average force monotonically increased as the standard stiffness value increased across the sessions. We also found positive correlation between the Weber fraction and the palpation speed in the sessions tested with 1.00 and 1.25 N/mm standard stiffness. This correlation suggested that higher speed is related to lower sensitivity in discrimination of stiffness differences for these two standard stiffness values. Our results are applicable to tasks involving stiffness discrimination between multiple objects.
(*Sim Healthcare* 5:000–000, 2010)

Key Words: Stiffness discrimination, Compliance, Haptics, Weber fraction, Palpation, Palpation velocity, Palpation force.

The usage of simulation systems is proceeding toward becoming an important part of medical education. These simulators for training of medical professionals (doctors, nurses, physical therapists, veterinarians, etc.) are beneficial for several different reasons, some of them are as follows:

- Patient safety is not a concern because the trainees do not deal with real patients, providing them an opportunity to safely learn from their mistakes.
- The simulated medical problems or cases are repeatable, ie, the same situation can be worked on several times until the desired results are achieved.
- Speed and accuracy of any task may be enhanced by repetition with feedback on performance (such as palpation, taking pulses, and insertion of tubes).

From the Interdisciplinary Institute for Neuromusculoskeletal Research, Ohio University, Athens, OH.
AQ: K

Supported by the Osteopathic Heritage Foundation of Columbus, OH.

Reprints: Robert L. Williams II, PhD, Department of Mechanical Engineering, 259 Stocker Center, Ohio University, Athens, OH 45701-2979 (e-mail: williar4@ohio.edu).
AQ: L

DOI: 10.1097/SIH.0b013e3181e9e783

<article-id pub-id-type="doi">10.1097/SIH.0b013e3181e9e783</article-id> • <fpage></fpage> • <lpage></lpage>

- Simulations may become powerful assessment tools because every trainee may be exposed to the same situation, and their performance may be measured qualitatively and quantitatively to be evaluated later.

The simulation systems can be in virtual reality (computer generated), augmented reality (virtual reality is reinforced by means of real accessories), or in the form of a mannequin that can imitate the symptoms of a real human. The addition of haptics (sense of touch) into virtual/augmented reality simulators increases their capabilities beyond visual and audio feedback alone by enabling the users to "feel" the virtual environment.

We are involved with development of haptic simulations for palpatory diagnosis. Palpation, an economical and effective first line of medical diagnosis used in many fields of healthcare, plays an important role in osteopathic, allopathic, and veterinary medicine. It is fast and inexpensive, but limited patient (human or animal) volume and uneven mix of pathologies make training of professionals in these areas difficult. For instance, training of osteopathic medical students on palpation is usually performed in laboratories with students practicing on each other. These settings do not provide the typical population (age and physical condition) that these students will diagnose and/or treat. Therefore, we developed the virtual haptic back (VHB) as a training tool for medical students.[1–3] It is a simulation of contours and tissue textures of a human back that is presented graphically and haptically. Students use haptic devices to "feel" the VHB and identify the dysfunctional region. The dysfunctional region is simulated as increased stiffness compared with the background stiffness of the palpable portion of the entire back.

The VHB is the only human back simulation that is actively being used in palpation training of osteopathic medicine students. E-Pelvis is another example of a palpation simulator.[4,5] It is an electronic mannequin, which enables users to see on a computer screen where in the pelvis they touch during training and the pressure they apply to those touch points. The availability of objective performance data (the applied pressure and number of times a critical point with a sensor is touched) makes it a powerful assessment tool as well. There are also examples of palpation simulators in veterinary medicine. The Bovine Rectal Palpation Simulator is used to teach veterinary medicine students to identify fertility problems and diagnose pregnancy.[6] Another example is the simulator developed for feline abdominal palpation training.[7]

It is apparent that palpation simulators are becoming more common and applicable as haptics technology is integrated into these simulations. The ability to discriminate stiffness differences is a very important and frequently used part of palpation. It is used to detect problems such as muscles in spasm, lumps in breasts, testes, and abdomens. For instance, the existence of a lump in breast tissue is initially identified by palpation before any further investigation is performed. Therefore, any simulation that is designed for diagnosis through palpation would, directly or indirectly, involve training of individuals to increase their sensitivity for stiffness discrimination. However, as in any sensory modality, humans are subject to certain limits in terms of discriminating the minimum stiffness difference between two or more objects. These limits, varying noticeably among individuals, may be the results of limited amount of feedback received from cutaneous and proprioceptive receptors and methods (high/low force, high/low speed, etc.) used during the palpation process. Simulations can be useful tools to help users improve their sensitivity. Users can also be taught and can practice the methods that will help them improve their limits (ie, higher sensitivity to discriminate subtle changes).

The Weber fraction is used to quantify the limits, ie, the minimum difference that a person can discriminate, for different sensory modalities. The Weber fraction is defined to be the ratio of this minimum difference to the standard intensity of the stimulus in a sensory modality. In the area of haptics, there have been a number of studies to find the limits of human perception of stiffness (or compliance, the inverse of stiffness) differences. DeGersem[8] used a PHANToM 1.5 haptic interface and found Weber fractions between 0.08 and 0.12. The support point of the arm for this study was the elbow. Jones and Hunter[9] reported a Weber fraction of 0.23 for stiffness discrimination by using a contralateral limb matching procedure.[7] With supported elbows, the subjects matched the stiffness of two motors (one matching and one reference) connected to their wrists. Tan et al[10] reported mean Weber fractions for stiffness of 0.08 for fixed and 0.22 for roving displacement. In their study, the subjects squeezed two plates along a linear track by using the thumb and the index finger. Howell et al[1] reported a Weber fraction of 0.11 for compliance detection after eight practice sessions (before practice, the Weber fraction was 0.40) with the VHB, a simulation of the feel of a human back. In this study, subjects were asked to identify randomly set regions with a higher stiffness value (representing altered tissue texture), using the shoulder as their support point. One of the differences among these previous studies is the varying support point of the arm or finger during palpation.

The accuracy of the stiffness discrimination between two objects in virtual and real-life situations may be affected by various measures such as the existence of tactile and/or kinesthetic feedback, visual feedback, stiffness value of the surfaces, and the speed of exploration. It is important to identify these effects to be able to develop better virtual environments. Srinivasan and LaMotte[11] analyzed the effect of the tactile and kinesthetic feedback and found that, for deformable objects, kinesthetic feedback alone is not sufficient to discriminate stiffness, whereas tactile feedback alone is sufficient. Srinivasan et al[12] showed the importance of visual dominance in their related article by using two virtual springs and providing the subjects with independent visual and kinesthetic clues. The authors asked the subjects to identify the stiffer spring and changed the relationship between the visual deformation and actual deformation on each spring throughout the trials. Their results showed a complete dominance of the visual feedback over the kinesthetic feedback.

In this article, we study the effect of surface stiffness on stiffness discrimination. We also measure and analyze palpation forces and speeds during haptic exploration. We hypoth-

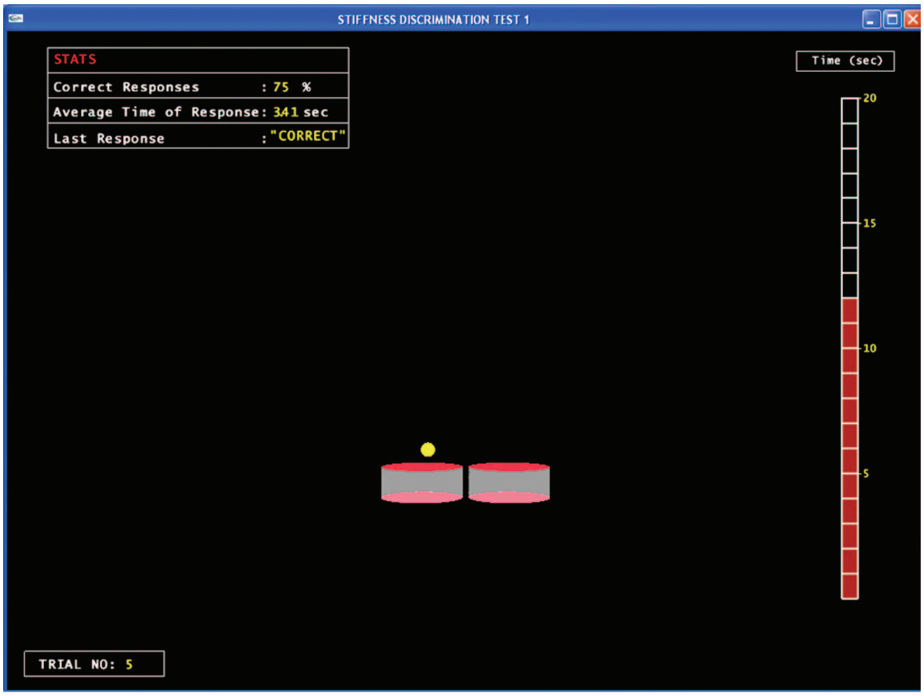

Figure 1. Visual scene of the experiment.

esize that the forces applied to detect the stiffness of an object in consideration increases as the stiffness increases.

The main objectives of this study were (1) to design and perform an experiment to test the ability of individuals to discriminate stiffness differences; (2) to analyze the palpation forces during palpation; and (3) to analyze palpation speed during haptic exploration. Our stiffness detection results compare well with previously published results, but our analysis of palpation forces and speed is a new result. We also discuss the possible contributions of this study to medical simulations.

METHODS

Experimental Setup

The experiments were run on a 2.8 GHz dual Pentium PC with 1 GB RAM and an NVIDIA Quadro 4XGL video adapter. A PHANToM Omni haptic interface displayed the stiffnesses to the subjects. The graphical interface was written using Microsoft Visual C++ and the OpenGL graphic library. The haptic effects were implemented by using the SensAble OpenHaptics Toolkit.

F1 As shown in Figure 1, two computer-generated virtual cylinders were created, whose top surfaces were haptically explored by the subjects. These surfaces did not visually deform because that would introduce another variable because of the complete dominance of the visual feedback over the kinesthetic.[11] One of these two cylinders always had the standard stiffness value (standard side) for that specific experimental session, and the stiffness value of the remaining cylinder (comparison side) was adjusted according to the subject's performance (correct or incorrect identification of the stiffer side) throughout the experiment. The surface with the standard stiffness value was always stiffer than the one with the comparison stiffness value. Which side, left or right, was the standard side varied randomly.

A virtual gap between the two computer-generated virtual cylinders prevented the subjects from detecting the change in stiffness by sliding the finger from one surface to the other. This gap forced the subjects to explore the two stiffness stimuli independently.

The stylus of PHANToM Omni haptic device was modified to be able to obtain stiffness values up to 1.25 N/mm. The design of the stylus enabled users to apply a force four times higher than the maximum force feedback capability of the Omni (3.3 N instantaneous according to the device manual). This was accomplished by moving the point of application of the force away from the Omni's original pivot (the distal end of the second arm) to the finger stylus shown in Figure 2. The F2 modification permitted the arms of the Omni haptic interface to be nearly orthogonal for most of the virtual surfaces manipulation because the device has the highest stiffness in the *y* direction (up-down motion of the subject's finger) in that configuration. The subjects were able to adjust the size of the finger holder by means of an adjustable strap for comfort during the experimentation.

Figure 3 shows an overview of the experimental setup with F3,AQ:I subject, PC display, and modified PHANToM Omni haptic interface.

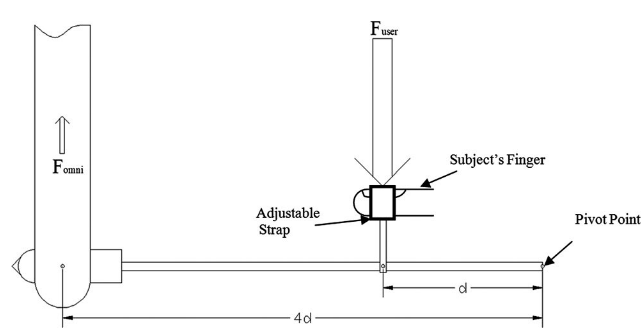

Figure 2. Diagram of the Omni-modified stylus.

Figure 3. Experimental setup.

Subjects

Thirteen adult subjects (eight women and five men, all osteopathic medical students) participated in the experiment. They had no known neuromuscular abnormalities. Eight of the subjects had prior experience with haptics (force and touch feedback from virtual reality) by participating in at least one of the studies conducted by the Virtual Haptic Back Laboratory of the Interdisciplinary Institute for Neuromusculoskeletal Research.[1–3,13] Ohio University IRB approval was obtained for this experiment, and all participating subjects signed an informed consent form.

Procedure

The experiment consisted of four different sessions to accommodate testing over four different standard stiffness values (0.25, 0.50, 1.00, and 1.25 N/mm). Each session was composed of a number of trials. The exact number of trials depended on the subjects' performance. The subjects completed all four sessions in 2 weeks, and they were not allowed to perform more than one session per day. The subjects were given a $20 gift certificate and a letter of participation for their medical school files.

The subjects were presented with two spatially separated stimuli that they probed actively (Fig. 1). The subjects put their index fingers in the custom-made stylus and were asked to identify the stiffer side of the two virtual surfaces. They were allowed to feel the stiffness of the surfaces as long as they wanted to until the 20-second time limit ended. Forces reflected to the subject were calculated using Hooke law, multiplying the depth of compression of the virtual object by the stiffness at that location. Of the two stimuli, one represented the standard stiffness side, whereas the other was the comparison stiffness side. The standard stimulus, which was constant throughout a session, was randomized to be either on the left or on the right side.

A two-alternative forced-choice procedure was used on each trial.[14] The subjects had two alternatives to choose from (right or left cylinder), and they had to select either one to be able to proceed to the next trial. They received feedback indicating whether they correctly identified the stiffer side, the running percentage of their correct responses, and the average time of response. The average time of response was provided to encourage the subjects to remain within a certain time limit (20 seconds for each trial), so that they could finish the session without becoming bored or fatigued.

To determine the subjects' stiffness discrimination threshold, we used an adaptive method. A threshold is defined to be the stimulus (in this case stiffness) difference that can be discriminated at a target level of performance. Adaptive methods do not require the a priori determination of the sequence of stimulus levels, and the threshold is found by systematically varying these stimulus levels during the experiment. There are four questions that need to be answered during the design of an adaptive method: (1) How should we decide to end testing at the current stimulus level and shift to a new one?, (2) Once we decide to end testing at the current stimulus level, what level should it be changed to?, (3) How should we end the experiment?, and (4) How do we calculate the estimated threshold? We answer these questions in the following paragraphs.

We used the Wald rule to decide when to shift to a new level of stimulus[14] (see Appendix A for more information). Once the decision to change to the next stimulus level is made, depending on whether the stimulus intensity (stiffness difference) was too high or too low, we need to decide the stimulus intensity of the next level. For that purpose, we change (increase or decrease) the previous stimulus level by one step. The size of the step, namely the increment or decrement of the stiffness difference, is determined by using parameter estimation by sequential testing (PEST,[14] see Appendix B for more information). PEST rules generally use decreasing steps. They allow quick recovery from obviously incorrect decisions. If a fixed step size were used, the recovery from these incorrect decisions would take more time making it even more difficult for the participant to concentrate on the task due to frustration and tiredness.

The experimental sessions started off with the smallest standard stiffness (0.25 N/mm) and continued in ascending order. The experimental sessions ended when the subject experienced seven reversals. A reversal is a step in the opposite direction from the previous step, for example, a stiffness decrease after a stiffness increase or vice versa. The average of the final four reversals determined the threshold or the just-noticeable difference for each subject. These just-noticeable differences were used to calculate the individual Weber fractions (*W*) of each subject for the four standard stiffness values using the equation: $\mathbf{W} = \frac{\textbf{JND}}{\text{standard stiffness}}$.[15]

The palpation velocity vectors and forces were recorded directly from the haptic device by using the predefined functions of the manufacturer's OpenHaptics Toolkit. Data were gathered when the sphere-shaped cursor (Fig. 1) and the surface of the virtual cylinders were in contact. The cursor contacted the surface at a single point, and this contact was frictionless. The recorded force and velocity values were multiplied and divided by four, respectively, to account for the lever arm action of the modified stylus (Fig. 2). The palpation speed is the magnitude of the velocity vectors. For each trial, the data were separated into standard side (surface with standard stiffness value) and comparison side (surface with com-

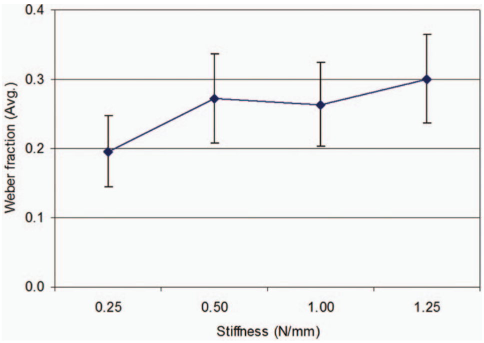

Figure 4. Weber fractions vs. standard stiffness values for all 13 subjects (mean ± standard error).

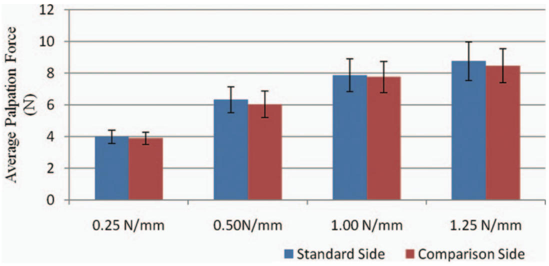

Figure 5. Average palpation forces applied by subjects on the standard and comparison fitness sides across four sessions (mean ± standard error).

parison stiffness value), and the average of the values was calculated. Average palpation speed and forces in this context are the grand average of these values for all the trials in a single session.

Statistical Analysis

The gender and haptic experience differences in Weber fraction values were analyzed using a multivariate analysis of variance. The four dependent variables used were the four Weber fraction values corresponding to the 0.25, 0.50, 1.00, and 1.25 N/mm standard stiffness values. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicolinearity, with no serious violations noted. For post hoc analysis, paired *t* tests

AQ: C with the Bonferroni correction for repeated test were used.

The correlation between the average palpation speed and Weber fraction, and total time spent by the subjects for each session and their corresponding Weber fraction values were analyzed using the Pearson product-moment correlation coefficient.

RESULTS

Weber Fraction

The average Weber fractions for the four standard stiffness

F4 values are shown in Figure 4. Four standard stiffness values (0.25, 0.50, 1.00, and 1.25 N/mm) were tested in four different sessions, and the average Weber fractions results were 0.20, 0.27, 0.26, and 0.30, respectively.

An inspection of the average Weber fraction values for male and female subjects indicated that men had slightly higher Weber fraction values than women, but the differences were not statistically significant ($P > 0.05$ for all stiffness levels tested). There were five haptically experienced and eight haptically inexperienced participants. The experienced users had slightly smaller average Weber fractions than the inexperienced ones except at the 0.25 N/mm standard stiffness value. These differences, however, were not statistically significant ($P > 0.05$ for all stiffness levels tested).

Palpation Force

The average palpation forces applied by the subjects on the comparison and the standard stiffness sides for all sessions

F5 are shown in Figure 5. Even though the palpation forces are close to each other within each session, they increase as the standard stiffness value increases. Repeated-measures analysis of variance revealed that when the average force exerted on the comparison side was compared with the average forces exerted on the standard side, there was no significant difference in any of the sessions. There were significant differences in the average force applied to standard and comparison sides between all sessions ($P < 0.05$) except between the sessions with 1.00 N/mm and 1.25 N/mm standard stiffness. As it can be seen from Figure 5, the subjects monotonically increased their average palpation force across the sessions as they proceeded from the smallest of the four standard stiffness values (0.25 N/mm) to the largest (1.25 N/mm).

The average forces applied to the standard side by individual subjects is shown in Figure 6. It is seen that the subjects in F6
general tend to increase their average forces across the sessions with increasing standard stiffness.

We also looked at whether users varied their palpation force during each session by breaking all trials into four quartiles (Fig. 7) based on the time sequence of the trial: ie, the F7
first 25% of trials performed in a session were grouped in the first quartile; the second 25% were grouped in the second quartile, etc. Repeated-measures analysis of variance revealed that there was no significant difference between the average force applied on standard and comparison sides in any quartile of any session. It also revealed that there was no significant difference in the average forces applied on standard or comparison sides across the quartiles of all trials in any session.

Palpation Speed

The average palpation speed of the subjects while probing the comparison and standard stiffness sides are shown in Figure 8. We found that there was no significant difference F8
between the average palpation speed of the subjects while probing the standard side and the comparison side.

There were positive correlations between the average palpation speed and average Weber fraction in sessions with 1.00 and 1.25 N/mm standard stiffness values. This correlation suggested that high Weber fraction (Table 1) is associated T1
with high average palpation speed at these two standard stiffness values. AQ: D

Figure 9 shows the subjects' Weber fractions versus aver- F9,AQ:J
age palpation speed (average of the palpation speed for all trials, for both standard and comparison sides). Again, from Figure 9, the subjects who have higher palpation speed mostly have higher Weber fraction values ($R^2 = 0.6512$).

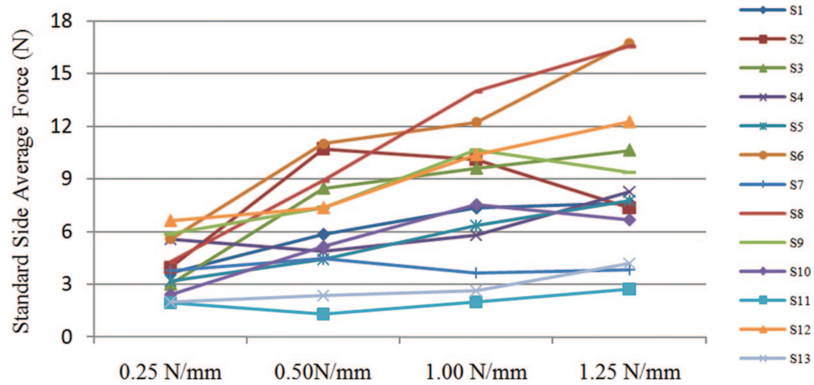

Figure 6. Average palpation forces applied by individual subjects on the standard side.

DISCUSSION

Weber Fraction

In this study, we measured the stiffness-detection Weber fractions of 13 subjects and analyzed the associated applied forces and the palpation speed. Our average experimental Weber fractions were 0.20, 0.27, 0.26, and 0.30, for the four standard stiffness values 0.25, 0.50, 1.00, and 1.25 N/mm, respectively. Lower Weber fraction indicates a higher sensitivity for stiffness difference discrimination.

Palpation Force

The results indicated that, at each standard stiffness level, the average of the forces applied by the subjects remained stationary across all four quartiles of all trials. In Choi et al,[16] it was shown that the stiffness of the surface played a significant role in determining the force level while they were detecting surface topography by haptically exploring two surfaces with different stiffness values and same height. Similarly, we showed that the average palpation force applied by the subjects remained constant within a session and monotonically increased as the standard stiffness value increased across the sessions. These results prove our hypothesis, which predicted that the applied forces would increase with stiffness. In our study, subjects were asked to discriminate an existing stiffness difference between two virtual computer-generated surfaces, whereas in study by Choi et al, subjects were instructed to probe a surface for 15 seconds to discriminate height of two surfaces (with no need of discriminating the stiffness).

Palpation Speed

We found that palpation speed and Weber fractions are positively correlated, ie, that higher speed is related to lower sensitivity to stiffness differences, in the sessions with 1.00 and 1.25 N/mm standard stiffness. This finding raises the question: Is there a range of palpation speeds while exploring objects with different stiffness values within which the users are protected from the degrading sensitivity? We cannot answer this question with this study because it needs to be explored further with a study in which velocity is the controlled independent variable. The question is whether higher palpation speed results in lower sensitivity in stiffness discrimination and/or whether lower discriminatory sensitivity in a subject (related to other factors) results in a compensatory increase in his or her palpation speed (or whether both effects can be present and interact). The answer to this question is important for virtual reality and real-life applications. For instance, there are two major palpation techniques that are taught to osteopathic medical students to diagnose medical problems by touch. The first technique is to palpate a body part with slow finger motions to identify any somatic dysfunctions (Glossary of Osteopathic Terminology, 2006, defines a somatic dysfunction as: "Impaired or altered function of related components of the somatic (body framework)

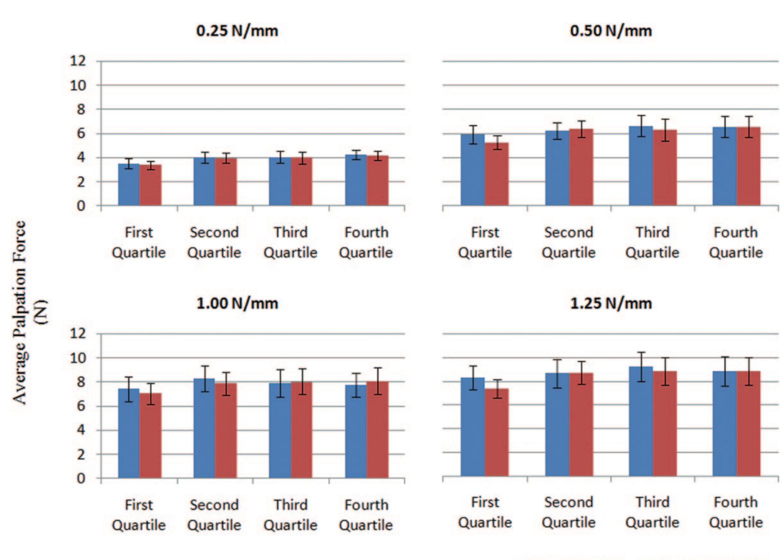

Figure 7. Average palpation forces applied by subjects on the standard and comparison stiffness sides in quarterly subsets of the total number of trials (mean ± standard error).

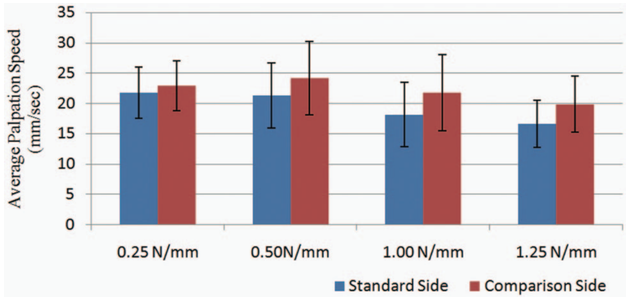

Figure 8. Average palpation speed by subjects on the standard and comparison stiffness sides across four sessions (mean ± standard error).

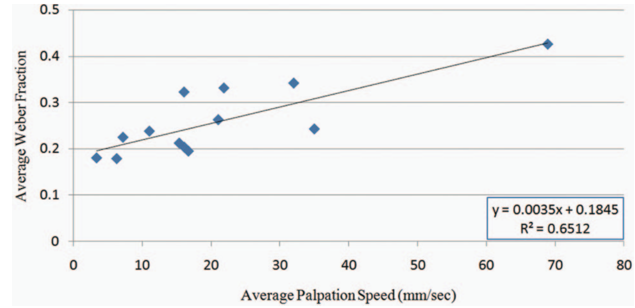

Figure 9. ●●●.

system: skeletal, arthrodial and myofascial structures, and their related vascular, lymphatic, and neural elements.") (mostly tissue texture changes), whereas in the second, the
AQ: E fingers are used percussively. Both techniques use the tissue texture to identify acute or chronic dysfunctions. In acute dysfunctions, tissue is edematous and collects fluids from vessels and chemical reactions in the tissue. In chronic dysfunctions, on the other hand, tissue can be doughy, stringy, and shows increased resistance. The percussion technique requires the utilization of higher palpation speeds, when compared with slow motion. If, in fact, there is a speed limit of palpation for an individual after which the sensitivity to discriminate stiffness decreases, then this outcome could be used in training of medical students in terms of choosing the best technique of learning and teaching palpation.

Previous Research

T2 Table 2 presents our findings (Weber fractions) in comparison with the results of previous stiffness discrimination studies. Our results (average Weber fractions: 0.20, 0.27) are comparable with the results of Tendick et al[17] (average Weber fractions: 0.17, 0.30) for the same nominal stiffness values (0.25, 0.50 N/mm) because the wrist support point is closer to the metacarpophalangeal joint than the elbow or the shoulder. The resolution of these two joints was also found to be the same in a study by Clark et al.[18] In their article, the authors presented a method to assess the exactness of position sensing, the target resolution method. This method estimates the maximum number of targets that can be resolved with no errors within a specified range. Their data indicate that the target resolution increases as the joint considered becomes more proximal and the metacarpophalangeal and the wrist joint have the same target resolution. DeGersem[8] found lower Weber fractions than this study for a similar standard stiffness value range, indicating that humans are more sensitive at discriminating stiffnesses when the elbow is the support point. It is important to consider the variability in the hardware used for all the aforementioned studies. There is not a study, to the authors' knowledge, that uses the same haptic device, range of standard stiffness values, or method to prove that the amount of proprioceptive feedback and/or the number of joints involved during palpation are significantly important to the stiffness discrimination task.

One concern of this experiment was the possibility of the subjects getting frustrated because each session took between 5 and 35 minutes, depending on their performance throughout that specific session. In psychophysical experiments, excessive completion time can negatively affect the performance of the subject. Therefore, it decreases the convergence rate of the algorithm used to estimate the threshold because the subjects can easily get frustrated by repetitively performing the same task. A few subjects did report some level of fatigue and frustration but mentioned that it did not obstruct their ability to discriminate the stiffness differences between the two virtual objects during the experiment. Supporting their subjective feedback, we found that the time they spent for each session was not positively correlated with the Weber fractions calculated for each subject for any of the four sessions ($P > 0.05$). This suggested that fatigue did not play a role as to degrading the discrimination performance of the subjects during any of the sessions.

Table 1. Correlation Between the Average Palpation Speed and the Average Weber Fractions

	Weber Fraction			
	Session 1 (0.25 N/mm)*	Session 2 (0.50 N/mm)*	Session 3 (1.00 N/mm)*	Session 4 (1.25 N/mm)*
Average palpation speed on the standard side				
Pearson correlation	0.342	0.514	0.596	0.649
P (two tailed)	0.253	0.073	0.032†	0.016†
Average palpation speed on the comparison side				
Pearson correlation	0.317	0.515	0.651	0.628
P (two tailed)	0.291	0.072	0.016†	0.021†
Average of palpation speed				
Pearson correlation	0.330	0.515	0.627	0.640
P (two tailed)	0.271	0.072	0.022†	0.018†

*The standard stiffness value in that session.

†Statistical significance, $P < 0.05$.

Table 2. Summary of Studies on Stiffness Discrimination

Source	Weber Fraction	Support Point	Stiffness (N/mm)	Note
Williams et al.[13]	0.09–0.21	Shoulder	≈0.4	PHANToM 3.0
Jones and Hunter[9]	0.19–0.28	Elbow	0–6.4	Contralateral limb matching (no use of haptic device)
DeGersem[8]	0.08–0.12	Elbow	0.3, 0.6, 0.9, 1.2	PHANToM 1.5
Tendick et al.[17]	0.13, 0.17, 0.30	Wrist	0.125, 0.25, 0.50	3 DOF PHANToM (wrist supported by a sliding track)
Current article	0.20, 0.27, 0.26, 0.30	Metacarpophalangeal joint	0.25, 0.50, 1.00, 1.25	Modified PHANToM Omni

Implications on Medical Simulations

In the latest version of the VHB, we continuously control the amount of force medical students apply during palpation. The main reason for that is, based on the feedback of expert osteopathic doctors, the appreciation of dysfunctional areas can be impaired by the application of more force than necessary to find the problem. In other words, palpatory information from superficial soft tissues can be lost with application of high forces. To teach that principle to the students training on the VHB, the intensity of the problematic area (increased stiffness compared with the rest of the back) decreases as the students increase their force and finally diminishes as they reach a predetermined amount of force. This predetermined force is currently based on the measurements that were done with force transducers during palpatory examinations. The testing software used in this study can be used to define that threshold force for any palpation simulator (involving stiffness discrimination) by using experts as subjects and analyzing their force data.

The complete method of the testing software that we designed for this study can also be used in training and performance assessment of users who make use of palpation for diagnosis and/or treatment. The Weber fractions and time to completion provided at the end of each session can be used as measures of improvement and performance of the users in discriminating stiffness differences. Some of the other modalities that could be used in testing or training of medical professionals using the same methodology can be discrimination of angle, magnitude of force, etc. In this direction, for instance, we are developing a family of modules including this stiffness discrimination experiment along with several other basic skill trainers. These modules are well underway at our institute and they are in evaluation by physicians at another medical institution.

CONCLUSION

For the four standard stiffness values investigated, 0.25, 0.50, 1.00, and 1.25 N/mm, the resulting average stiffness-discrimination Weber fractions were 0.20, 0.27, 0.26, and 0.30, respectively. The average of the forces applied by the subjects was constant within a single session (with a single standard stiffness value). This average force monotonically increased as the standard stiffness value increased across the sessions. We also found positive correlation between the Weber fraction and the palpation speed in the sessions tested with 1.00 and 1.25 N/mm standard stiffness values. This correlation suggested that higher speed is related to lower sensitivity in discrimination of stiffness differences for these two standard stiffness values. Our results are applicable to tasks involving stiffness discrimination between multiple objects.

APPENDIX A—THE WALD RULE

The Wald rule determines whether a change should be made in the stimulus intensity presented to a test subject based on the subject's sequential performance in the test. Specifically, if the number of correct responses by a subject is outside of a prespecified range, the rule directs that the stimulus intensity be changed as detailed below.

The Wald rule determines whether the current level of a variable (in our case, the stiffness difference) results in an event proportion less or greater than the target probability, $\mathbf{P_t}$. When a new level is started, the running count of number of correct answers, $\mathbf{N(C)}$, and the total number of trials, $\mathbf{T}$, are recorded. After completion of each trial within this level, permissible lower and upper bounds are defined on $\mathbf{N(C)}$. If $\mathbf{N(C)}$ is above the upper bound, then the level (the stiffness difference) is "too high," and if $\mathbf{N(C)}$ is below the lower bound, the level (the stiffness difference) is "too low." In the case where $\mathbf{N(C)}$ is between the lower and upper bounds, the trials for the same level continue until a decision is made. If the current level being tested is exactly the same as the target level (the level at which the proportion correct is equal to $\mathbf{P_t}$), then the expected number of correct trials after $\mathbf{T}$ is as follows:

$$E[\mathbf{N(C)}] = \mathbf{P_t} \cdot \mathbf{T} \tag{1}$$

The upper and lower bounds are found by adding and subtracting a constant $\mathbf{W}$, respectively.

$$\mathbf{N_{UB}} = E[\mathbf{N(C)}] + \mathbf{W} \tag{2}$$

$$\mathbf{N_{LB}} = E[\mathbf{N(C)}] - \mathbf{W} \tag{3}$$

where $\mathbf{N_{UB}}$ and $\mathbf{N_{LB}}$ are the upper and lower boundaries on $\mathbf{N(C)}$, respectively. Small values of $\mathbf{W}$ will result in quick but not very powerful decisions, whereas a large $\mathbf{W}$ will require

Table 3. Upper and Lower Boundaries for the Wald Rule ($W = 1.0$)

Trial Number	E[N(C)]	N_{UB}	N_{LB}
1	0.75	1.75	−0.25
2	1.50	2.50	0.50
3	2.25	3.25	1.25
4	3.00	4.00	2.00
5	3.75	4.75	2.75
⋮	⋮	⋮	⋮

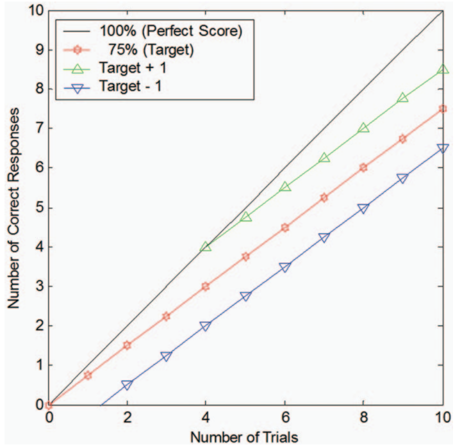

Figure 10. Trials correct vs. total trials (adapted from Macmillan and Creelman, 1991).

more trials but result in relatively more powerful decisions, ie, for larger values of **W**, the number of trials it takes to make a decision will increase, but the decision will be more accurate in terms of estimating whether the current level of stiffness difference is in fact too high or too low.

In this study, we defined the target probability $\mathbf{P_t}$ as 0.75, halfway between chance for a two-alternative trial (0.5) and certainty (1.0), and **W** as 1.0. Table 3 shows a sample (first
T3
five trials only) calculation of the upper and lower boundaries to make a decision with the chosen **W** value. According to the table, two consecutive incorrect answers at the beginning of a new level will be below the lower boundary ($\mathbf{N(C)}$ = 0.0 $\leq N_{LB}$ = 0.50), and the decision will be that the level (stimulus intensity) is "too low." On the other hand, four consecutive correct answers will reach the upper boundary ($\mathbf{N(C)}$ = 4.0 $\geq N_{LB}$ = 4.0) and the decision will be that the level is "too high."

According to the Wald rule, the actual number of trials needed to make a decision varies depending on the perfor-
F10
mance. Figure 10 illustrates the Wald rule used in this study by plotting the number of correct responses against the total number of trials at a single level for a series of trials.

In Figure 10, if the subject performance at the current level is above the upper limit (Target +1), then it is said that this level of stimulus intensity, in our case the stiffness difference, is "too high." On the other hand, if the performance drops below the lower limit (Target −1), the stimulus intensity is decided to be "too low." A decision to change the stiffness difference is not made as long as the subject's performance lies in the area encompassed by these two limits. Once the decision is made, the stimulus intensity is changed in a particular direction (increase or decrease) by using the step size as determined by the PEST rules (Appendix B).

APPENDIX B—PEST RULES

In this study, we wanted to determine the minimum stiffness difference that can be discriminated by individuals. We used the PEST,[14] which is an adaptive procedure for rapid and efficient psychophysical testing.[19] In adaptive methods, such as the PEST, the level (intensity of the stimulus) to be tested at any given trial is determined by a portion of the

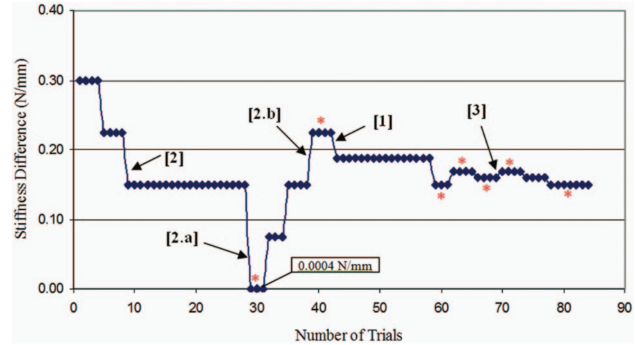

Figure 11. Demonstration of the PEST rules (actual subject data, reversals are shown with an asterisk).

history of the run. These methods aim to make measurements at levels near the target level (the minimum stiffness difference discriminated by an individual). After each trial, the current level is tested with the Wald rule (Appendix A) to change the stimulus level in a particular direction (increase or decrease). The next stimulus level is adjusted by incrementing or decrementing by a step size. The methods that use fixed step sizes take a long time to recover from possible incorrect decisions. The PEST, however, uses variable step sizes to address this problem. The PEST rules generally use decreasing step sizes but use increasing ones to rapidly recover from an incorrect decision. The PEST rules are basically designed to adjust the step size. These rules are[12] as follows:

1. If two blocks of trials at different levels resulted in opposite answers, then the target level is most probably between these two levels. Therefore, after each reversal (a reversal is a step in the opposite direction from the previous step, for example, a stiffness decrease following a stiffness increase or vice versa.) the step size is halved ([1] in Fig. 11), until: F11,AQ:F
2. The step size remains the same for two consecutive changes in the same direction ([2]). However, if the Wald rule gives the same result several times in a row, then the step size is increased more to shift toward the target level rapidly. Therefore, the step size is doubled when the third change is in the same direction as the previous 2 ([2.a]). An exception occurs when a reversal happens right after doubling the step size. The step size will then double when there are three consecutive changes in the same direction instead of 2 ([2.b]).
3. The step size remains the same when it reaches a prespecified minimum value ([3]).

ACKNOWLEDGMENTS

The authors thank David Noyes, PhD, for building the Omni™ haptic interface modified stylus for this study. They also thank Joseph Samosky, PhD, University of Pittsburgh, Pittsburgh, PA, for his suggestions and comments, which enhanced the quality of this paper during the review process.

REFERENCES

1. Howell JN, Williams RL II, Conatser RR Jr, Burns JM, Eland DC. Training for palpatory diagnosis on the virtual haptic back: performance improvement and user evaluations. *J Am Osteopathic Assoc* 2008;108:29–36.

2. Howell JN, Williams RL II, Conatser RR Jr, Burns JM, Eland DC. *The Virtual Haptic Back (VHB): Detection of Compliance Differences*. Paper #06DHM-27, SAE Digital Human Modeling Conference, Lyon, France; July 4–6, 2006.

3. Howell JN, Williams RL II, Conatser RR Jr, Burns JM, Eland DC. *The Virtual Haptic Back (VHB): A Virtual Reality Simulation of the Human Back for Palpatory Diagnostic Training*. Paper #05DHM-63, SAE Digital Human Modeling Conference, Iowa City, IA; June 14–16, 2005.

4. Pugh CM, Heinrichs WL, Dev P, Srivastava S, Krummel TM. Use of a mechanical simulator to assess pelvic examination skills. *JAMA* 2001; 286:1021–1023.

AQ: G

5. Pugh CM, Youngblood P. Development and validation of assessment measures for a newly developed physical examination simulator. *J Am Med Inform Assoc* 2002;9:448–460.

6. Baillie S, Crossan A, Brewster S, Mellor D, Reid S. Validation of a bovine rectal palpation simulator for training veterinary students. *Stud Health Technol Inform* 2005;111:33–36.

7. Parkes R, Forrest N, Baillie S. A mixed reality simulator for feline abdominal palpation training in veterinary medicine. *Stud Health Technol Inform* 2009;142:244–246.

8. DeGersem G. *Kinesthetic Feedback and Enhanced Sensitivity in Robotic Endoscopic Telesurgery*. Ph.D. Thesis, Katholieke Universiteit Leuven, Belgium; 2005.

9. Jones LA, Hunter IW. A perceptual analysis of stiffness. *Exp Brain Res* 1990;79:150–156.

10. Tan HZ, Durlach NI, Beauregard GL, Srinivasan MA. Manual discrimination of compliance using active pinch grasp: the roles of force and work cues. *Percept Psychophys* 1995;57:495–510.

11. Srinivasan MA, LaMotte RH. Tactual discrimination of softness. *J Neurophysiol* 1995;73:88–101.

12. Srinivasan MA, Beauregard GL, Brock DL. The impact of visual information on the haptic perception of stiffness in virtual environments. *Proc ASME Dyn Syst Control Division* 1996;58:555–559.

13. Williams RL II, Howell JN, Conatser RR Jr, Burns JM, Eland DC. *Palpatory Training on the Virtual Haptic Back Improves Detection of Compliance Difference*. Eurohaptics, Paris, France; 2006.

14. Macmillan NA, Creelman CD. *Detection Theory: A User's Guide*. Cambridge: Cambridge University Press; 1991. AQ: H

15. Gescheider GA. *Psychophysics: the Fundamentals*. 3rd ed. Lawrence Erlbaum Associates, Mahwah, NJ; 1997:3.

16. Choi S, Walker L, Tan HZ, Crittenden S, Reifenberger R. Force constancy and its effect on haptic perception of virtual surfaces. *ACM Trans Appl Percept* 2005;2:89–105.

17. Tendick F, Cavusoglu MC, Dhruv N, Sherman A. *Maximizing the Sensation of Compliance in Teleoperative Surgery*. Proceedings of the Eighth International Symposium on Robotics With Applications, Maui, HI; June 11–16, 2000.

18. Clark FJ, Larwood KJ, Davis ME, Deffenbacher KA. A metric for assessing acuity in positioning joints and limbs. *Exp Brain Res* 1995;107:73–79.

19. Taylor MM, Creelman CD. PEST: Efficient estimates on probability functions. *J Acoust Soc Am* 1967;41:782–787.