Proceedings Paper:

Jamieson, ES, Chandler, JH, Culmer, PR et al. (3 more authors) (2015) Can Virtual Reality Trainers Improve the Compliance Discrimination Abilities of Trainee Surgeons? In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 25-29 Aug 2015, Milan, Italy. IEEE , pp. 466-469. ISBN 978-1-4244-9270-1

https://doi.org/10.1109/EMBC.2015.7318400

Can Virtual Reality Trainers Improve the Compliance Discrimination Abilities of Trainee Surgeons?*

Earle S. Jamieson, James H. Chandler, Peter R. Culmer, Michael Manogue, Mark Mon-Williams, Richard M. Wilkie

Abstract— The assessment of tissue compliance using a handheld tool is an important skill in medical areas such as laparoscopic and dental surgery. The increasing prevalence of virtual reality devices raises the question of whether we can exploit these systems to accelerate the training of compliance discrimination in trainee surgeons. We used a haptic feedback device and stylus to assess the abilities of naïve participants to detect compliance differences with and without knowledge of results (KR) (groups 1 and 2), as well as the abilities of participants who had undergone repetitive training over several days (group 3). Kinematic analyses were carried out to objectively measure the probing action. Untrained participants had poor detection thresholds (mean just noticeable difference, JND = 33%), and we found no effect of KR (provided after each trial) on performance (mean JND = 35%). Intensive training dramatically improved group performance (mean JND = 12%). Probing action (in particular, slower movement execution) was associated with better detection thresholds, but training did not lead to systematic changes in probing behaviour. These findings set a benchmark for training systems that act to increase perceptual sensitivity and guide the learner toward optimal movement strategies to improve discrimination.

I. Introduction

Humans perform skilful interactions with objects using a combination of visual and haptic information [1]. There are few domains where skillful and successful interactions of this type are as critical as surgery. For example, it is common for dental surgeons to use instruments such as a blunt dental probe to confirm the health of a tooth, as extensive tooth decay (dental caries) alters the compliance of the tooth's structure [2]. Whilst the use of probing techniques to obtain information about dental health is common, it is also reported as being difficult to teach students and can take a long time to master [3]. Tissue palpation during laparoscopic surgery is another technique that is difficult to perform [4]. This raises the question of whether virtual reality (VR) simulators could be used to improve the compliance discrimination abilities of surgical students. To our knowledge, only one study [5] has shown that repetitive training can lead to improvements in compliance discrimination performance. Unfortunately, there was no investigation into whether these improvements were due to: (i) systematic changes in probing strategy; (ii) increased perceptual sensitivity to compliance cues or (iii) a combination of (i) and (ii).

Compliance perception requires the monitoring of force and displacement information during interactions with an object [6]–[9]. This is achieved by using a combination of cues from haptic (cutaneous and kinesthetic) receptors located in the skin, muscles and joints, as well as vision [10]–[13]. The quality and quantity of information obtained via these sources seems to depend on the exploratory procedure used. For instance, during direct finger interactions, rich compliance cues about the geometrical changes of an object's surface are available. Conversely, when using a tool to interact with a surface the compliance response can be perceived at the fingerpad via shear or normal forces through the tool, but geometrical information is not present [14]. The type of movement used to interact with a compliant object also appears to affect performance. Findings from previous work on tool-based interactions revealed that slow movements and low forces result in the best discrimination performance [15], whilst others [16] have found that tapping is more effective than pressing. For direct interactions, finger force and velocity seem to be adjusted strategically to the expected compliance [17], implying that these variables play a significant role in our perception of compliance. These findings indicate that optimal probing strategies might exist. However, there is a need for an investigation into whether discrimination performance can be improved with practice, and if so, what gives rise to improved performance.

We developed a system to measure the just noticeable difference (JND) of simulated tissue compliance in adults, along with the kinematics of their movements. We had three training groups: Group 1 consisted of novice participants that received no explicit feedback (knowledge of results, KR) after each trial (this mimics normal conditions where trainees must become attuned to intrinsic visual-motor feedback in the absence of performance feedback). Group 2 were given KR after each trial (a feature possible within virtual reality training systems). Group 3 received identical training to Group 2, except that they had intensive training (12 sessions) over a week to determine the extent to which compliance discrimination can be improved through training with performance feedback. We also examined the relationship between probing strategy and performance, and whether there were systematic changes in probing action during training.

* Research supported by EPSRC

E. S. Jamieson is with the School of Mechanical Engineering and the School of Psychology, The University of Leeds, Leeds, United Kingdom, LS2 9JT (email: el07esj@leeds.ac.uk).

J. H. Chandler is with the School of Mechanical Engineering, University of Leeds, Leeds, United Kingdom, LS2 9JT (email: mn07jhc@leeds.ac.uk) .

P. R. Culmer is with the School of Mechanical Engineering, The University of Leeds, Leeds, United Kingdom, LS2 9JT (email: P.R.Culmer@leeds.ac.uk).

M. Manogue is with the School of Dentistry, The University of Leeds, Leeds, United Kingdom, LS2 9JT (email: M.Manogue@leeds.ac.uk).

M. Mon-Williams is with the School of Psychology, The University of Leeds, Leeds, United Kingdom, LS2 9JT (email: M.Mon-Williams@leeds.ac.uk).

R. M. Wilkie is with the School of Psychology, The University of Leeds, Leeds, United Kingdom, LS2 9JT (email: R.M.wilkie@leeds.ac.uk).

II. METHODS

A. Materials and experimental design

Custom software was developed to enable communication between a haptic device with a stylus end-effector (PHANToM Omni, SensAble Technologies) and a bespoke user interface developed in a graphical development environment (LabVIEW, National Instruments). This enabled the simulation of different modulus values with corresponding visual information provided via a computer monitor. As illustrated in Figure 1(a), Participants were seated directly in front of the monitor and instructed to rest their right arm and wrist on supports. Three groups were tested: novice - no KR (Group 1), novice – KR (Group 2), and trained - KR (Group 3). First, between-subjects tests were carried out to assess the effects of KR on the performance of novices. A repeated measured design was then used to test the learning effect of the trained subjects.

B. Task Configuration

Participants actively probed two virtual samples, one after the other, and then judged which was the least compliant (stiffest). An adaptive staircase algorithm, the PEST [18], [19], was employed to generate the stimuli on a trial by trial basis. A correct or incorrect judgement on one trial was used to adjust the stimulus (stiffness) on the next. The threshold obtained at the end of each set of 100 trials was identified as the JND value (this number of trials was selected based on pilot trials to allow successful convergence on a JND value, whilst providing the same number of training trials to all participants). A constant 'baseline' stiffness of 0.075 N/mm was used for one sample in each trial. The other sample could be given an 'offset' stiffness value ranging from 0 to 0.1 N/mm above the baseline. These stiffness values are similar to the properties of soft human tissue such as liver [20]. The order of appearance (first or second) of the baseline sample was randomised. The first trial of each session contained a 'baseline + offset' sample located in the middle of the offset range at 0.125 N/mm.

C. Participants

The research was approved and conducted under the guidelines established by The University of Leeds Research Ethics Committee. All participants were Psychology or Engineering undergraduate or PhD students. Nineteen participants were recruited and randomly allocated to one of the three groups: Group 1 consisted of seven participants (3 male, 4 female, 22-28 years, M = 24 years) who performed one session. Group 2 consisted of six participants (3 male, 3 female, 21-27 years, M = 23.7 years) who also performed one session but with KR: a green 'tick' was displayed for a correct response and a red 'cross' for an incorrect response. Group 3 consisted of six participants (4 male, 2 female, 21 to 28 years, M = 22.5 years). The procedure for Group 3 was identical to that of group 2 (KR was provided), except that each participant completed twelve sessions over four days (three sessions per day - morning, noon and late afternoon).

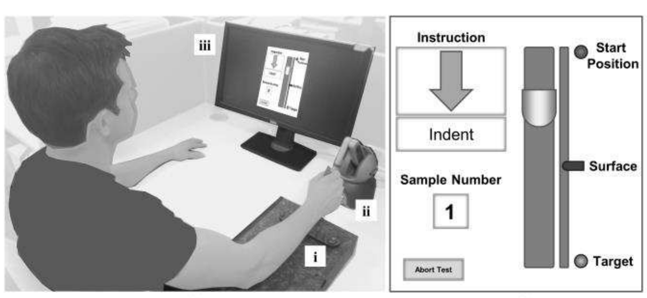

Figure 1. (a) Diagram of the experimental setup, showing i) the arm and wrist supports, ii) the haptic feedback device, and iii) the graphical display. (b) A close-up of the graphical display; the instruction panel informs the participant to 'Indent', 'Await Instruction', or 'Move to the start position'; 'Sample number' denotes which sample (1 or 2) is displayed. The position of the stylus tip is updated at a rate of 30 Hz

D. Procedure

On each trial, participants used the haptic device to indent each sample of simulated tissue before identifying which of the two was perceived as the stiffest. For each sample, participants were required to move the stylus downwards from a "start" position past the "surface" until the "target" indentation depth was reached (see Figure 1(b)). An auditory tone was used to indicate when participants could start their movements. A higher pitched tone then indicated when they had reached the target indentation depth. Forces were generated as a function of indentation depth to simulate an object with fixed compliance and a cursor on the monitor indicated the position of the stylus. Participants vocally identified the stiffest sample by saying "One" or "Two". The experimenter electronically recorded their response and the next trial was then presented. No explicit instructions were given to participants regarding what movement characteristics to use (e.g. speed and acceleration were unconstrained). However, all participants were required to hold the stylus in a standard precision grip and they were informed of the vertical probing movements that would be required. Movement was free in all three dimensions. All participants received a practice run of 25 trials prior to the start of the experiment to familiarise them with the task and to ensure they understood the instructions. Each session consisted of 100 trials and lasted approximately 15 minutes.

E. Kinematic analysis

We explored the movement kinematics of the probing actions over time in order to determine the relationship between probing strategy and performance. Figure 2 illustrates the virtual probe and its relative position to the surface during the probing action. For each sample, time and probe position in the vertical axis were recorded at 100 Hz. To objectively assess probing strategy, two kinematic metrics were calculated for each sample: strike velocity (V_S) and

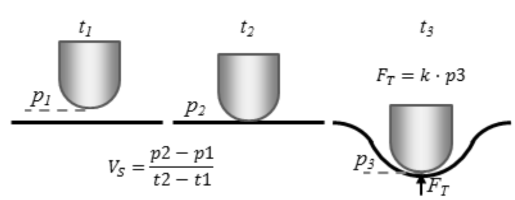

Figure 2. Illustration of the virtual probe at different stages of indentation, showing i) the probe at 2 mm prior to contacting the virtual surface, (p_1 = -2 mm), ii) the instant at which the probe comes into contact with the surface (p_2 = 0 mm), and iii) the deepest indentation (p3) during the probing action, when FT is calculated. The variables t_X and p_X denote the time and relative position of the probe to the surface at each probing stage, respectively

terminal force (F_T). V_S was calculated as the average velocity of the probe as it traveled from position p_1 to position p_2. It was not influenced by interactions with the sample, making it useful for assessing movement behaviour independent of sample compliance (which varied from sample to sample). F_T was calculated as the product of the sample's stiffness coefficient, k, and the position at maximum indentation, p_3. The time points (t_X) were specified as the time elapsed from the start of the trial until each corresponding p_X location was reached. Time measurements were accurate to within 0.5 milliseconds. In order to measure probing consistency, we also calculated the absolute difference in V_S (ΔV_S) and F_T (ΔF_T) between the first and second indentations. Note that due to the relationship between k and p_3, ΔF_T should only be assessed for trials where identical samples are presented. This is to avoid the effect of natural variations in F_T if one or more other variables were controlled by the subject during probing of surfaces with different compliance.

III. Results

The compliance JND values and mean kinematic metrics for all trials obtained for Groups 1, 2 and Session 1 of Group 3 are shown in Table 1. The JND of one participant in Group 1 did not fall below 100% and so they were excluded due to failure to perform the task. An independent-samples t-test revealed no significant difference between the JNDs of groups 1 and 2 ($t(10) = 1.145$, $p = .279$, $r = .34$), suggesting the possibility that KR had no significant effect on performance. There were also no reliable differences between Group 2 and Session 1 of Group 3 ($t(10) = -.31$, $p = .76$, $r = .098$). To assess longer term training on JND performance we examined changes across each session for Group 3 (Figure 3). A repeated-measures ANOVA showed a significant main effect of Session on JND ($F(11,55) = 3.15$, $p = .002$, $\eta_P^2 = .39$), with participants gradually improving compliance sensitivity over time from 26.6% to a best value of 12.1% in session 7 (though values then drifted to 16%, possibly reflecting exploration of different perceptual strategies).

To determine whether changes in probing strategy could explain the improvements in compliance sensitivity we examined the kinematic metrics (Figure 4a). There was a significant correlation between V_S and JND whereby improved sensitivity to compliance differences was associated with slower probing velocity ($r = .476$, 95% BCa CI [.254, .674], $p<.001$). A similar relationship was also found between ΔV_S and JND, ($r = .526$, 95% BCa CI [.309, .711], $p<.001$), and between F_T and JND, ($r = .484$, 95% BCa CI [.256, .695], $p<.001$). Two participants (P17 and P18) employed particularly fast probing actions, but if they were excluded from the correlations the pattern of results did not change. We also observed a strong relationship between V_S and F_T, ($r = .909$, 95% BCa CI [.807, .953]), and V_S and ΔV_S, ($r = .865$, 95% BCa CI [.744, .931]) ($ps < .001$), consistent with slower movements leading to lower terminal forces and greater consistency of probing. To determine whether probing behavior altered across sessions, we performed a repeated-measures ANOVA on each kinematic measure. These analyses revealed no significant effect of Session on V_S ($F(11,55) = 0.97$, $p = .49$, $\eta_P^2 = .16$), ΔV_S ($F(11,55) = 1.38$, $p = .21$, $\eta_P^2 = .22$) or F_T ($F(11,55) = 1.59$, $p = .13$, $\eta_P^2 = .24$). We also performed repeated-measures ANOVAs on the consistency measures, which failed to demonstrate a significant main effect of Session on ΔV_S ($F(11,55) = 0.76$, $p = .68$, $\eta_P^2 = .13$), or ΔF_T ($F(11,55) = .78$, $p = .66$, $\eta_P^2 = .13$). These results suggest that the group's JND improvements over the training period were not due to changes in probing strategy captured by V_S and F_T, nor by increased consistency in probing actions (as measured by ΔV_S and ΔF_T). It seems therefore that JND improvements were achieved by tuning into the appropriate perceptual information [21].

Finally we assessed the extent to which individual differences explained the relationship between probing behaviours and JND performance (Figure 4b). Three

TABLE I. THE MEAN COMPLIANCE JND, V_S, ΔV_S, AND F_T ACROSS ALL TRIALS FOR GROUP 1, 2 AND SESSION 1 OF GROUP 3. 'M' AND 'SD' ARE THE MEAN AND STANDARD DEVIATION OF THE MEAN, RESPECTIVELY

Group no.	JND (%)		V_S (mm/s)		ΔV_S (mm/s)		F_T (N)	
	M	SD	M	SD	M	SD	M	SD
1	35.0	9.4	80.5	47.6	1.21	0.23	27.6	15.9
2	34.8	16.9	60.6	36.7	1.09	0.22	20.6	8.1
3(1)	26.5	5.2	68.8	53.1	1.11	0.33	22.6	13.3

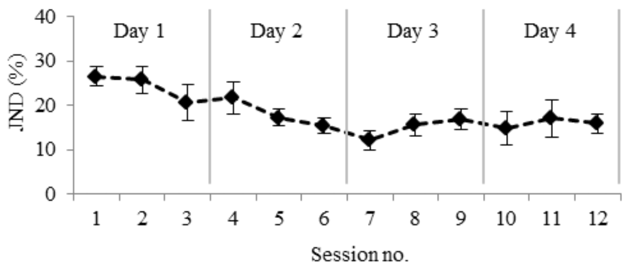

Figure 3. Mean JND obtained at each training session for group 3. Error bars indicate the standard error of the mean

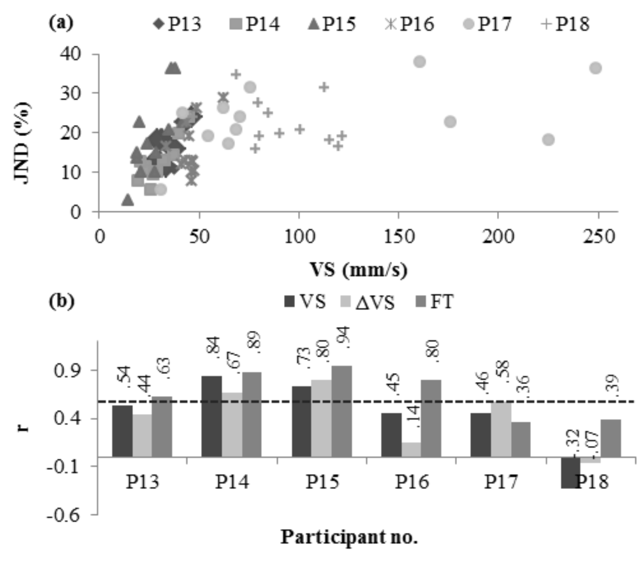

Figure 4. (a) Mean Strike Velocity (V_S) against JNDfor each training session and each participant, (b) correlation coefficients for each participant obtained for the mean V_S, ΔV_S and FT against JND obtained during all training sessions. The horizontal line indicates the significance threshold ($r = .497$), above which the correlation becomes significant at the $p = .05$ level

individuals match the group level analysis – a clear relationship between all three kinematic metrics and JND performance. However there were also three individuals who displayed a weaker relationship between probing kinematics and JND and these individuals were those that exhibited the highest mean V_S, ΔV_S and F_T across all sessions. With a group size of six it is difficult to draw firm conclusions about these differences, but it seems that there is a non-linear relationship between probing kinematics and performance, and that different probing strategies may be adopted across individuals. For instance, the magnitude of V_S had a strong positive relationship with JND for P14, whereas for P18, there was a negative relationship suggesting that they may have tuned into information related to F_T (or another unmeasured variable). These differences could explain some of the inconsistencies in performance between individuals within Groups 1 and 2, but it is clear that further work is needed to determine the degree to which individual differences can explain the perceptual learning of compliance when probing.

IV. Conclusions

Training with feedback in a VR trainer over an extended period could be an effective method of improving performance in tasks requiring the discrimination of small compliance differences using a handheld tool. In addition to increased perceptual sensitivity, understanding how surgical trainees can be effectively guided towards optimum movement strategies to maximise their performance will be crucial. Further research is needed to identify when probing is performed within a 'sweet spot' that avoids the disadvantages of moving either too quickly or too slowly. It must be the case that very slow movements impair sensitivity because slow movements are generally more difficult to execute smoothly, which may have implications for sensory performance [22]. In contrast, very fast movements must limit the perceptual information available from the probing action. Understanding what training programmes could be effectively implemented to guide a surgeon towards these movement strategies will be critical for future applications. For this study we used a relatively low stiffness range that is comparable to soft tissues such as liver. However, the same methods described here could be applied to investigate compliance detection for different stiffness ranges, such as when indenting tooth structures with a periodontal probe (though a haptic device would be required that is capable of a larger force output than the one employed in the present study). Our findings have potential implications for robotic dental and surgical training systems, where trainees are given information and guidance during the execution of practice tasks [23]. In order to successfully translate our research into clinical settings, however, it is important to evaluate the compliance range and mechanical properties of tissues involved, as well as the active and perceptive abilities required by the trainee to complete their tasks. Possessing an underlying theoretical framework which describes the relationship between these factors could lead to a more robust understanding of how suitable training strategies should be implemented, and of how they can be transferred to surgical disciplines that span from dentistry to laparoscopic surgery.

References

[1] M. Mon-Williams, J. P. Wann, M. Jenkinson, and K. Rushton, "Synaesthesia in the normal limb," Proc. R. Soc. Lond. B Biol. Sci., vol. 264, no. 1384, pp. 1007–1010, Jul. 1997.

[2] R. H. Selwitz, A. I. Ismail, and N. B. Pitts, "Dental caries," The Lancet, vol. 369, no. 9555, pp. 51–59, Jan. 2007.

[3] S. D. Drucker, L. E. Prieto, and D. W. K. Kao, "Periodontal Probing Calibration in an Academic Setting," J. Dent. Educ., vol. 76, no. 11, pp. 1466–1473, Nov. 2012.

[4] P. Culmer, J. Barrie, R. Hewson, M. Levesley, M. Mon-Williams, D. Jayne, and A. Neville, "Reviewing the technological challenges associated with the development of a laparoscopic palpation device," Int. J. Med. Robot., vol. 8, no. 2, pp. 146–159, 2012.

[5] K. Teodorescu, S. Bouchigny, and M. Korman, "Training Haptic Stiffness Discrimination Time Course of Learning With or Without Visual Information and Knowledge of Results," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 55, no. 4, pp. 830–840, Jan. 2013.

[6] S. Choi, L. Walker, H. Z. Tan, S. Crittenden, and R. Reifenberger, "Force constancy and its effect on haptic perception of virtual surfaces," ACM Trans. Appl. Percept., vol. 2, no. 2, pp. 89–105, 2005.

[7] Kaim and K. Drewing, "Exploratory strategies in haptic softness discrimination are tuned to achieve high levels of task performance," Haptics IEEE Trans. On, vol. 4, no. 4, pp. 242–252, 2011.

[8] Tan, N. I. Durlach, G. L. Beauregard, and M. A. Srinivasan, "Manual discrimination of compliance using active pinch grasp: The roles of force and work cues," Percept. Psychophys., vol. 57, no. 4, pp. 495–510, Jun. 1995.

[9] Tan, N. I. Durlach, Y. Shao, and M. Wei, "Manual resolution of compliance when work and force cues are minimized," Adv. Robot. Mechatron. Haptic Interfaces, vol. 49, pp. 99–104, 1993.

[10] M. Kuschel, M. Di Luca, M. Buss, and R. L. Klatzky, "Combination and Integration in the Perception of Visual-Haptic Compliance Information," IEEE Trans. Haptics, vol. 3, no. 4, pp. 234–244, Oct. 2010.

[11] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, "Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review," Psychon. Bull. Rev., pp. 1–33, Nov. 2012.

[12] M. A. Srinivasan, G. L. Beauregard, and D. L. Brock, "The impact of visual information on the haptic perception of stiffness in virtual environments," in ASME Winter Annual Meeting, 1996, vol. 58, pp. 555–559.

[13] W. M. B. Tiest and A. M. L. Kappers, "Cues for Haptic Perception of Compliance," IEEE Trans. Haptics, vol. 2, no. 4, pp. 189–199, 2009.

[14] R. M. Friedman, K. D. Hester, B. G. Green, and R. H. LaMotte, "Magnitude estimation of softness," Exp. Brain Res., vol. 191, no. 2, pp. 133–142, Nov. 2008.

[15] E. Karadogan, R. L. Williams, J. N. Howell, and R. R. Conatser, "A Stiffness Discrimination Experiment Including Analysis of Palpation Forces and Velocities:," Simul. Healthc. J. Soc. Simul. Healthc., vol. 5, no. 5, pp. 279–288, Oct. 2010.

[16] R. H. LaMotte, "Softness Discrimination With a Tool," J. Neurophysiol., vol. 83, no. 4, pp. 1777–1786, Apr. 2000.

[17] L. R. Kaim and K. Drewing, "Finger force of exploratory movements is adapted to the compliance of deformable objects," in World Haptics 2009., 2009, pp. 565–569.

[18] H. Lieberman and A. Pentland, "Microcomputer-based estimation of psychophysical thresholds: The Best PEST," Behav. Res. Methods, vol. 14, no. 1, pp. 21–25, 1982.

[19] M. Leek, "Adaptive procedures in psychophysical research," Atten. Percept. Psychophys., vol. 63, no. 8, pp. 1279–1292, 2001.

[20] S. Mueller and L. Sandrin, "Liver stiffness: a novel parameter for the diagnosis of liver disease," Hepatic Med. Evid. Res., vol. 2, pp. 49–67, May 2010.

[21] J. Tresilian, Sensorimotor Control and Learning: An introduction to the behavioral neuroscience of action. London: Palgrave Macmillan, 2012.

[22] H. Nagasaki, "Asymmetric velocity and acceleration profiles of human arm movements," Exp. Brain Res., vol. 74, no. 2, pp. 319–326, Jan. 1989.

[23] S. Khan and L. L. Cabanilla, "Periodontal probing depth measurement: a review," Compend. Contin. Educ. Dent., vol. 30, no. 1, pp. 12–14, 16, 18–21; quiz 22, 36, Feb. 2009.