A new computer application for teaching sexual history taking to medical students: innovation and evaluation in the UfaceME™ program

Background: We developed an app (UfaceME) which simultaneously allows patient and clinician to be viewed on a split screen, then replayed and rated on a series of semantic differential scales and the ratings of the clinician of their own response to the interview, and the patient’s response, displayed on a graph.

Method: We evaluated the app with trials with 14 medical student volunteers who alternated the roles of patient and clinician, using 2 randomly allocated sexual case histories. Semantic differentials for 10 adjectives were examined pre- and post-exercise. There were significant differences in 8 of the 10 adjectives. A focus group with the 14 subjects was also carried out and transcribed. The exercise was repeated with all 175 first-year medical students taking a sexual history.

Results: Qualitative data indicated themes of being surprised and educated by non-verbal and verbal responses; seeing how the patient rated their responses; rating and watching discomfort; differences in clinician/patient perceptions and response; and the advantage of feedback. Quantitative data comparing self vs “patient” showed low correlations between perceptions of being “at ease”, moderate for appearing “distracted”, and high for appearing “engaged”.

Conclusion: UfaceME was easily understood and used, and the opportunity to replay and rate performance on key semantic scales, and watch and assess verbal and non-verbal performance, including patient rating, provided valuable insight.

Keywords: sexual history, interview skills app, medical students, training feedback

Introduction

Teaching sexual history taking to medical students is recognized as a crucial area of clinical medicine, but unfortunately in many medical schools is limited to didactic material or not taught at all.1,2 Because sexual history taking is an area that both patients and practitioners consider challenging, embarrassing, or threatening, it is crucial to develop skills based on practice and competencies rather than simply provision of information.

The conduct of sexual history taking has been covered in detail in a number of monographs and reviews.3 However, training in sexual history taking and studies which describe and compare methods of training are not extensive.4 Given the move to measuring competencies in medical and other health care professional training, it is important to develop skills in sexual history training that are measurable and which enhance the performance, rather than just the knowledge, of medical students.
Althof and colleagues\textsuperscript{5} have described the crucial domains that need to be considered in assessing the competence of medical students in taking sexual histories. These include strengthening the therapeutic alliance and enhancing the patient–clinician relationship; increasing confidence in being able to take a sexual history; and developing empathy and rapport. They note that patients may perceive that practitioners are reluctant, disinterested, or unskilled in sexual problem management, and that practitioners are disinclined to ask about sexual issues due to their negative attitudes about sexual issues, time constraints, unrealistic fears about offending the patient, deficits in communication skills, reimbursement concerns, the lack of available or approved treatments, and a growing gap between sexual medicine and the clinical skills of practicing clinicians.

Typically, role plays have been used, both scripted (to reduce discomfort and invasive privacy issues) and improvised, with medical students as a means of getting interaction within a large group.\textsuperscript{6} Such an approach, while being a common method, still suffers from a “blind leading the blind” approach, unless the participants have simultaneous adequate training in sexual health and supervision during the process. Getting closer to measuring actual competencies, using simulated patients (SPs) in objective standardized clinical examinations (OSCEs), has the advantage of using skilled actors and having the SPs provide feedback on the spot to the trainee practitioners, although higher costs are involved. Feedback, both visual (through replay of interview videos) and through verbal feedback from SPs, is key to enhancing performance.\textsuperscript{7} Interestingly, correlations between marks in sexual health stations compared with psychiatry and medical history taking were, respectively, 0.22 and 0.14, suggesting that markedly different skills may be involved beyond empathy and communication. Haist et al\textsuperscript{8} reported that following a SP-based 4 hrs workshop, participants scored one standard deviation higher than non-participants on an OSCE (Objective Structured Clinical Examination).

Other research into the efficacy of sexual history taking has suggested that computer-assisted self-interview (CASI) with face-to-face (FTF) sexual history-taking in a clinical setting produces largely similar responses, with clinicians being comfortable with the method.\textsuperscript{9} However, Kurth et al\textsuperscript{10} compared ACASI (Audio-CASI) with a clinician interview in the US, and found that ACASI produced significantly more reports of same-sex behavior, oral sex, transactional sex, and drug use, but fewer reports of STI symptoms. In a review of teaching sexual history taking, Coverdale, Balon, and Roberts\textsuperscript{3} described 11 trials, of medical students and residents, and judged only one trial to be of very high quality, concluding that evidence supported interactive workshops over didactic presentations. They concluded that (p.1590) “Clearly, the academic medicine community must do more to address this neglected curricular and competence domain in medical training.”

We report on the development, introduction, and preliminary evaluation of a new technology for video/audio recording medical student sexual history taking and feedback and its preliminary qualitative evaluation, the UfaceME\textsuperscript{TM} computer application (“app”).

Method

The computer application

Focused on FTF interaction, the UfaceME method (summarized in Figure 1) is a dynamic four-step relational learning tool which (1) mutually engages two participants in a conversation or interview, (2) video recording (based on a prism camera which places both side-by-side on a laptop screen), (3) rating their interpersonal experiences of their conversation as they view video playback, and (4) reviewing immediate feedback of their separate viewpoints as presented in graphic displays which link their covert cognitions to overt behaviors.

The video recording of a sexual history occurs with a patient history that is provided, and the students divide into the “clinician” and the “patient” (and may then exchange roles with a new case history). On review of the video of the interview, the UfaceME reporting structure is preset and consists of a series of statements presented at regular intervals, to which participants respond on a five-point Likert scale using handheld numeric keypads (eg, “You are empathic” or “I am empathic”). This allows participants to ascribe quantitative values to qualitative data, making it amenable to statistical analysis. A cursor can be moved to point on the graphed responses to watch the actual video at that point to see and review why there may have been discrepancies or high or low scores.

Each UfaceME conversation generates a personal database for each participant consisting of the video record and the quantitative values, all of which can be recalled for further study focused on relationship learning, supervision,
training, and research. Figure 1 summarizes the UFACEME app and Table 1 illustrates its setup. The UfaceME app is described in more detail at http://UfaceME.com/, and is protected by four US patents.

Pilot study
Our initial objective was to determine if the UfaceME was acceptable to medical students as a learning application, and to clarify if the adjectives used to evaluate the video interviews could be reduced in number to a few dimensions. Twelve second-year medical students (6 men, 6 women) at a large Midwestern state US medical school were recruited by class announcement invitation to participate in a trial and focus group to evaluate the UfaceME app. Students used the UfaceME app as both a “clinician” and a “patient”, using two provided case sexual histories (one male, one female) taking 8 mins each, and a focus group of 20–25 mins followed the exercise. Two investigators led the focus group and a third investigator took notes of participant responses. Participants were compensated with a $50 purchase card. The exercise was approved by the University of Minnesota IRB (Study no. 00002750) as exempt for the purposes of training design and evaluation. Nevertheless, participants provided informed consent and photo permissions, and participation was voluntary.

Full class exercise and evaluation
The full medical school Year 1 class (n=175) was randomized into dyads and took two sexual histories for 8 mins each as part of the Essentials of Clinical Management and Human Sexuality courses, position as “patient” and “clinician” also being randomized. Following each history taking, the video was replayed and the nine adjectives presented at adjectives and times:

- I feel at ease 0:00 mins
- You seem at ease 0:13
- Conversation was Appropriate 0:26
- I feel engaged 0:40
- You seem engaged 0:53
- Conversation was Empathic 1:06
- I feel distracted 1:20
- You seem distracted 1:33
- Conversation was On Track 1:46

(repeated every 2 mins). The resulting graphs were then displayed to the students and discussed for interview technique feedback by the tutor. Tutors were counseling graduate students or professionals. Students were not graded for the exercise, and it was presented as an opportunity to get feedback on sexual health interviewing skills. An example appears as Figure 2.

Measures
The pilot study adjectives on which the interviews were rated were initially selected based on the course objectives, including comfort, empathy, engagement, competence, or appropriateness of questioning, and for the focus groups, the liking and difficulty of the app process. Initial adjectives used are illustrated in Table 2. For the full class study, adjectives were selected based on the pilot study adjective factor analysis, described in the Analysis section below, which was carried out to reduce the number of adjectives to a few crucial dimensions (factors), and reduce question duplication to
a minimum. They consisted of three state adjectives (I feel at ease, You seem at ease, I feel engaged, You seem engaged, I feel distracted, You seem distracted) and three content-related measures (Appropriate, Empathic, On track) scored on a five-point Likert scale, using a handheld keypad.

Analysis
Quantitative data were analyzed by SPSS version 22 and by R 3.5. Comparisons between times (Time 1, Time 2) were computed by t-test between pairs, and between genders by t-test between independent groups. Factor analysis used Principal Axis extraction followed by Direct Oblimin rotation (Δ =0). Qualitative data were analyzed by separating comments into themes. All statistical tests were 2-tailed, with significance set at the 5% level.

Results
Pilot study
Factor analysis of the adjective scores in the Pilot Study (Table 2) indicated that there were three latent dimensions. These were (1) being engaged or distracted; (2) being at ease and seeing the conversation as appropriate; and (3) the interview questions being “on track”. Table 3 illustrates the changes pre- and post-training in taking a sexual history, with eight out of ten of the adjectives achieving statistical significance despite the small sample size.

Full class study
The first-year medical student class was 40% male and 60% female, with a mean age of 25.35 (SD=3.08). Data comparing the first and the second interview are presented in Table 4. Comparison of each student’s rating of the Likert scales at Time 1 (the first video) and Time 2 (the second video), where students were randomized into starting as “clinician” or “patient”, shows that on three of the nine scales (I feel engaged; Appropriate; and On track) there were significant differences between Time 1 and Time 2, all in the direction of increased scores over time. One other scale (I feel distracted) approached significance. On the first dyad, there were female–male mean differences for “I feel distracted” (1.99 vs 1.63, t=2.90, df=253, p= 0.01) and “You seem distracted”
(1.60 vs 1.39, $t=2.90$, $df=253$, $p=0.04$), with males reporting themselves and assessing interviewees as more distracted, but this gender difference had disappeared for both ratings by the second dyad (Time 2). Table 5 illustrates the mean ratings of “clinician” and “patient”: “clinicians” consistently rated themselves more poorly than the “patients” rated them.

Qualitative data from the pilot study focused on student response to the app and the process. Five themes emerged.

**Theme: seeing body language and hearing voice**

Relevant comments in this theme included: “Non-verbal feedback great”. “Helpful for me to see how I looked”. “Reading facial expressions was helpful and interesting”. “Seeing BOTH facial expressions simultaneously was great”. “How


Table 4 Mean on adjectival scales for first and second video experience in full class sexual history taking

<table>
<thead>
<tr>
<th>Variable</th>
<th>Video 1 mean</th>
<th>Video 2 mean</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel at ease</td>
<td>3.77</td>
<td>3.84</td>
<td>710</td>
<td>−0.85</td>
<td>0.22</td>
</tr>
<tr>
<td>You feel at ease</td>
<td>3.95</td>
<td>3.87</td>
<td>672</td>
<td>0.20</td>
<td>0.14</td>
</tr>
<tr>
<td>I feel engaged</td>
<td>4.37</td>
<td>4.49</td>
<td>644</td>
<td>−2.71</td>
<td>0.02</td>
</tr>
<tr>
<td>You seem engaged</td>
<td>4.37</td>
<td>4.39</td>
<td>633</td>
<td>−0.72</td>
<td>0.35</td>
</tr>
<tr>
<td>I feel distracted</td>
<td>4.13</td>
<td>4.26</td>
<td>607</td>
<td>2.06</td>
<td>0.07</td>
</tr>
<tr>
<td>You seem distracted</td>
<td>4.47</td>
<td>4.54</td>
<td>583</td>
<td>0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>Appropriate</td>
<td>4.50</td>
<td>4.61</td>
<td>642</td>
<td>−2.70</td>
<td>0.02</td>
</tr>
<tr>
<td>Empathic</td>
<td>4.00</td>
<td>4.08</td>
<td>613</td>
<td>−1.38</td>
<td>0.15</td>
</tr>
<tr>
<td>On track</td>
<td>4.12</td>
<td>4.36</td>
<td>591</td>
<td>−3.04</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Note: Bold p-values indicate statistically significant.

Table 5 Mean on adjectival scales for “patient” and “clinician” in full class sexual history taking

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patient mean</th>
<th>Clinician mean</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel at ease</td>
<td>4.16</td>
<td>3.79</td>
<td>1197</td>
<td>6.28</td>
<td>0.001</td>
</tr>
<tr>
<td>You feel at ease</td>
<td>4.36</td>
<td>3.90</td>
<td>1076</td>
<td>9.00</td>
<td>0.001</td>
</tr>
<tr>
<td>I feel engaged</td>
<td>4.65</td>
<td>4.42</td>
<td>984</td>
<td>6.04</td>
<td>0.001</td>
</tr>
<tr>
<td>You feel engaged</td>
<td>4.79</td>
<td>4.36</td>
<td>860</td>
<td>11.76</td>
<td>0.001</td>
</tr>
<tr>
<td>I feel distracted</td>
<td>1.71</td>
<td>1.79</td>
<td>1092</td>
<td>−1.31</td>
<td>0.10</td>
</tr>
<tr>
<td>You seem distracted</td>
<td>1.46</td>
<td>1.47</td>
<td>1052</td>
<td>−0.19</td>
<td>0.43</td>
</tr>
<tr>
<td>Appropriate</td>
<td>4.87</td>
<td>4.55</td>
<td>811</td>
<td>9.47</td>
<td>0.001</td>
</tr>
<tr>
<td>Empathic</td>
<td>4.70</td>
<td>4.03</td>
<td>828</td>
<td>15.78</td>
<td>0.001</td>
</tr>
<tr>
<td>On track</td>
<td>4.77</td>
<td>4.12</td>
<td>662</td>
<td>11.80</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Note: Bold p-values indicate statistically significant.

I felt and how I thought I felt and how my face looked was interesting”. “I used my hands too much”. “It was helpful to see the nonverbal signs and hear the TONE of voice”. “The tone betrays how you come across”. “Surprises included how I sounded”. “Great video technology to be able to speak face to face but also have video ‘head on’”. “Highlights were seeing my body language and hearing my linguistic choices”.

Theme: lack of realism of having the partner play a patient
Relevant comments in this theme included: “A real patient’s voice may be more authoritative”. “Could be useful for patients (or simulated patients) giving feedback”.

Theme: benefits of feedback and broader benefits
Relevant comments in this theme included: “Needs to be earlier in course”. “Has benefits in many other histories in other areas of medicine”. “Challenging in a good way”. “It does not feel forced or intrusive”. “You can be anxious about saying the wrong thing – it is helpful to get feedback”. “Getting feedback with tutor explaining was important” [for graphs]. “Wanted assessment and evaluation [not graded] by small group tutor”. “I could record in own time and earn points if it was in the course”. “Being able to track the similarities and differences in our feelings and responses during the conversations and being able to replay those nonverbal moments”. “Differentiating between uncomfortable discussions and feeling uncomfortable”. “Seeing the discrepancy between how I seem and how I feel”.

Theme: benefits of being able to flag moments on playback to illustrate graph points
Relevant comments in this theme included: “Benefits of flagging points and going back and discussing them”. “Being able to flag an awkward (or especially good) moment”. “Flagging capacity is key because it wasn’t always relevant to the moment”. “Being able to scroll through the video to identify moments on the graph”.

Theme: problems with technology
Relevant comments in this theme included: “Some confusion over responding and how to respond, especially ‘I feel’ and ‘you seem’”. “Confusion between clinician- and patient-scoring”. “Not super-intuitive”. “Critical to go back to the actual minute and it’s better with the video”. “Too many words to score, too many labels”.

Discussion
While feedback-informed counseling training (FIT) has been used and evaluated for some years, it typically involves the use of pencil and paper questionnaires and general symptom level routine outcome monitoring, and feedback with patients. The combination of video.
recording and immediate feedback in training, particularly in “difficult” interviewing areas, gives immediate feedback from the perspectives of the trainee and the simulated “patient”. The understanding that immediate and clear feedback is crucial to training is well established, but provision of almost instant feedback of visual and verbal feedback by both trainee and “patient” of themselves and each other on visual rating scales, along with split-screen playback, in a single package, is novel and powerful. It is especially important for training for “difficult” interviews such as sexual history taking, giving bad news, aggressive patients, and some psychiatric and criminological settings.

The data suggested strong student interest and enthusiasm in the UfaceME app for taking a sexual history, and the quantitative data suggested improvement over time on several scored dimensions of rating the experience. Randomization of the entire class (not a sample) into dyad, clinician or patient, and sexual history at Time 1 or Time 2 made it possible for some short-term comparisons of changes over time during the interviews.

The dimensions of the variables in the factor analysis in the Pilot Study suggest that what is being perceived is the level of engagement – or apparent interest – in the interview, along with a sense of being comfortable with the content area (being at ease and appropriate). The final dimension deals with progress through the clinical interview, being on track or appearing distracted. There has been little research looking at salient dimensions of perception in dyads in sexual history taking, and the present study is limited to medical students rather than clinician-patient dyads, making it difficult to generalize results to actual clinical dyad perceptions. Nevertheless, these data do suggest that within the dyad, level of engagement or distraction in the interview, the comfort level of the clinical interview dyad, and whether the interaction is seen as getting somewhere or moving into distractions are perceived as salient dimensions. Psychotherapy research has identified the dimensions of Genuineness, Accurate Empathy and Non-possessive Warmth as being causally related to psychotherapeutic success, and the present data suggest that the first latent dimension is a measure which includes empathy and sensitivity in the clinical dyad. This is perhaps not surprising given the nature of the interaction, which raises emotionally difficult and challenging subjects. The second dimension appears to measure mastery of the subject (the sexual history interview), while the third appears to relate to outcome – taking a history which is logical in sequence and arrives at a clarification or diagnosis. These may be used as dimensions for evaluating the process of learning to take a good sexual history, independent of the clinical content of the interaction.

Data from the Main Study with the full first-year class of over 170 medical students showed that even in the short duration of the study, comparing the first dyad and the second dyad, both randomly assigned, as were the two sexual histories, showed significant improvement in mean rating for three of the nine Likert scales: I feel engaged, Appropriate, and On track, with a fourth (I feel distracted) of borderline significance. Even in the short time between Dyad 1 and Dyad 2, there was clearly a significant improvement in ratings on the process of taking a sexual history (Appropriate and On track), and a reduction in feeling distracted. The data indicated that the “I feel distracted” rating was significantly higher for male students in the first dyad, but that the gender difference had disappeared by the second dyad. While comparisons with longer intervals between dyads, and between students who trained using the UfaceME app as opposed to the standard dyadic history taking without the app are needed to provide further assessment of the effectiveness of the app, these preliminary data suggest that there is a measurable increase in skills even after two applications. This does not seem to be due to increases in comfort, ease or empathy as these scales were not significantly different, but mainly to content and direction of the sexual history taking. However, in the absence of comparison with the standard (unassisted) role play which is the default method of teaching sexual history taking, this explanation remains speculative.

There was also a consistent effect of the “clinician” rating themselves lower than the “patient”. “Clinicians” were often reassured that they appeared to handle taking sexual histories better that they perceived themselves as doing, and that the “patient” also consistently perceived that. The importance of reinforcing confidence in difficult clinical interviews was commonly expressed by the medical students on seeing their data graphed.

Qualitative data, however, strongly confirm the indication of improvement. In the Pilot study with 12 students and 24 dyads, the five themes confirmed the students’ enthusiasm about the non-verbal feedback of watching their body language and mannerisms, and tone of voice. They also noted that the realism was somewhat limited by having a medical student play the patient, and the advantage that having a simulated patient would provide. A further theme noted the many benefits with history taking with difficult histories in general and the importance of getting such feedback early in the medical course.
Seeing the differences between how the “clinician” appeared (and was rated by the “patient”), and how they felt, was an important realization.

A further theme was general strong support for being able to ask difficult questions and get feedback, particularly with a tutor going over the video and highlighting points for reinforcement or alternative suggestions or approaches, eg, for awkward or good moments. The fact that points where there are differential assessments by the “clinician” and the “patient” can be instantly focused on by a click at that point on the ratings graph and jumping the video to that point was appreciated.

Finally, students noted some of the difficulty in using the technology which could be improved by making it more intuitive to use in a seamless way (which was done by modifying the app to make it more intuitive and user-friendly before the full class exercise).

For the Full Class exercise, data suggest that even between the first and second interview, there was improvement in “clinician” engagement, and in the two task-oriented Likert measures of the clinical conversation being “appropriate” in direction and “on track” (proceeding logically in taking the history). The fact that there were no increases in the measures of being at ease or empathic suggests that the improvement is not just a result of feeling more comfortable with the app. It is likely that such increases in comfort and non-judgmental interviewing style will take longer to develop.

It should be noted that the app was used in conjunction with a subsequent 20 hrs mandatory Sexual Health course which provided content, including patient panels offering consumer perspectives and dialog. Having an app which provides comprehensive training and feedback in the competency and comfort of a clinically relevant setting, and at the same time supportive content in the medical area and interaction with articulate patient panels, offers an ideal opportunity for addressing broadly sexual health education for medical students. The app fills an educational and training gap that has long been filled by “blind leading the blind” role playing without video replay and rating of the interview.

Nevertheless, there are limitations in this formative work. There were only two training episodes in the full class study, which lasted an hour, and a longer period of training is likely to be needed to demonstrate major changes. Further, a comparison with the standard training (FTF role play with occasional tutor feedback by interruption) would be a more accurate direct comparison.

**Conclusion**

The pilot study provided positive quantitative and qualitative feedback on the UfaceME sexual history taking app, and using the app on laptops in a live class training exercise demonstrated, in a total medical school class sample, that it was practical and successful to introduce to the curriculum and provided significant data even given the short duration of the exercise. The app appears to be a technologically and educationally superior alternative to current role playing approaches and provides videoed verbal and non-verbal feedback, an important educational principle for insight and improvement in skills training. The ratings by self and partner also provide feedback from a peer as well as an indication of accuracy of the trainee’s own perceptions. With a tutor to provide commentary, it appears to provide excellent feedback and training in verbal and nonverbal sexual history taking comfort and techniques. It could equally appropriately be used in other “difficult” interviews in medical, nursing, and other health settings with minimal modification of the rating adjectives to meet the objectives of training in interview skills.

**Acknowledgment**

This research was supported by a donation from Dr Duane Rost to the Joycelyn Elders Chair Discretionary Fund.

**Disclosure**

Dr Ayers, Mr Schmidt, Mr Bugbee, and Ms Knight are UfaceME employees and report conflict of interest. Dr James Ayers reports personal fees from Michael Ross, MD, PhD, outside the submitted work. In addition, Dr Ayers has a patent 9099010 issued, a patent 9262539 issued, a patent 9571753 issued, a patent 9832392 issued; and he is CEO and Co-Founder of UfaceME, Inc. Mr William Schmidt reports a patent 9099011 licensed to UfaceME, Inc., a patent 9262539 licensed to UfaceME, Inc., a patent 9571753 licensed to UfaceME, Inc., a patent 9832392 licensed to UfaceME, Inc.; Mr Thomas W. Bugbee reports he is the COO & CFO of UfaceME, Inc. which is assigned 4 US patents (noted above) regarding the UfaceME communications learning tool. Drs Ross, Muthyala and Newsstrom report no conflicts of interest in this work.

**References**


