Virtual Healthcare Interaction:
Papers from the AAAI Fall Symposium (FS-09-07)

Using Virtual Patients to Train Clinical Interviewing Skills

Barbara Hayes-Roth, Rami Saker, and Karen Amano

Lifelike Solutions, Inc.
54 Marianna Lane, Atherton, CA 94027
bella@lifelikesolutions.com

Abstract

Virtual patients are viewed as a cost-effective alternative to standardized patients for role-play training of clinical interviewing skills. However, training studies produce mixed results. Students give high ratings to practice with virtual patients and feel more self-confident, but they show little improvement in objective skills. This *confidence-competence gap* matches a common cognitive illusion, in which students overestimate the effectiveness of training that is too easy. We hypothesize that cost-effective training requires virtual patients that emphasize *functional* and *psychological fidelity* over *physical fidelity*. We discuss 12 design decisions aimed at cost-effective training and their application in virtual patients for practicing brief intervention in alcohol abuse. Our STAR Workshop includes 3 such patients and a virtual coach. A controlled experiment evaluated STAR and compared it to an easier E-Book and no-training Control. E-Book subjects displayed the illusion, giving high ratings to their training and self-confidence, but performing no better than Control subjects on skills. STAR subjects gave high ratings to their training and self-confidence and scored better higher than E-Book or Control subjects on skills. We invite other researchers to use the underlying Imp™ technology to build virtual patients for their own work.

Standardized vs. Virtual Patients

Patient-centered clinical interviewing can promote health, prevent disease, and manage chronic problems (Lein 07, Lyles 01, Miller 06). Role-play with standardized patients is the gold standard for training, but it is expensive and only moderately effective (Baer et al 2003, Miller 2006, Moral et al 2004, Rubel et al 2000). Medical educators call for improvements in training effectiveness (Lurie 2003, Miller et al 2006, Commonwealth 2002), with more practice on more cases, individualized feedback and coaching, and repeat practice to correct errors (Barney & Shea 2007, Clark 2003, Druckman & Bjork 1994). Using standardized patients, the cost of this would be prohibitive.

Virtual patients offer a promising alternative (Deladisma et al 2007, Deterding et al 2005, Hayes-Roth et al 2004, 2009, Johnsen et al 2007, Kenny et al 2007, Lok et al 2006, Orton et al 2008, Parson et al 2008, Raij et al 2007, Stevens et al 2006, Triola et al 2006). They scale to serve more students, reliably perform target protocols, allow systematic variation of cases, record comprehensive student data, and are available for practice anywhere, any time. Thus, they have the potential to reduce cost, increase access, and improve efficacy. However, despite compelling demonstrations, training results are mixed. Students praise virtual patients and feel more confident, but show little improvement in objective skills. In general, "there is no relationship between self-reported level of confidence and formally assessed performance" (Barnsley et al 2004, Miller et al 2006). Moreover, building confidence without competence might do more harm than good.

The Confidence-Competence Gap

This *confidence-competence gap* may reflect a common cognitive illusion: students overestimate the effectiveness of easy training tasks (Druckman & Bjork 1994). If virtual patient role-play is easy, students might mistakenly infer that their training is effective and their skills are strong.

As in all simulator training, practice effectiveness with virtual patients depends on *fidelity*, or resemblance to real patients. Limited resources preclude uniform high fidelity, but judicious resource allocation can produce cost-effective training. "The level of fidelity needed in simulators depends on the results of a task analysis of the specific skill to be taught. [This] can identify the critical elements in the skill that should be identical in the training and performance environments. Only when those factors are understood can resources be balanced against outcomes to design optimal training programs" (Druckman & Bjork 1994). Conversely: "If both the training system and operational setting share properties with regards to the objectives of training, other aspects of the training system could tolerate lower levels of fidelity without compromising effectiveness" (Alexander et al 2005).

We distinguish training benefits of 3 fidelity dimensions (Alexander et al 2005). With *physical fidelity*, which promotes training of perceptual-motor skills, a patient "looks, sounds, and feels like" a real patient. With *functional fidelity*, which promotes training of cognitive and interactive skills, a patient "acts like [a real patient] in reacting to the tasks executed by the trainee." With *psychological fidelity*, which promotes robust skills in the presence of psychological factors, the patient "replicates the psychological factors (i.e, stress, fear)."

Based on this analysis, we hypothesize that *most virtual patients fail to train interviewing skills effectively because they emphasize physical over functional fidelity*. High-fidelity physical elements may include life-size animation, realistic facial expressions and body language, eye contact, gaze control, gesture recognition, human voice, or speech understanding. Because these features induce a compelling sense of physical presence (Sanchez-Vives & Slater 2005), these patients might be effective for training clinical examination skills. Low-fidelity functional elements may include menu-based input, one-way reactive conversation, tightly programmed scripts, or role-play interruption for error correction. Because these features do not induce a compelling immersion in the interviewing task *per se* (Witemer & Singer 1994), these patients would not be effective for training clinical interviewing skills. Moreover, because they finesse the challenges of interviewing real patients, role-play with them is easy, inviting students to overestimate training effectiveness and their own skills.

Fidelity Decisions for Virtual Patients

Table 1. Fidelity Elements and Decisions

Fidelity Element	Fidelity Decision
High Functional Fidelity	
1. User input	Free-form natural language
2. Conversation	Mixed-initiative
3. Behavior	Adaptive to context
4. Expression	Normal variability
5. Reinforcement	Intrinsic
6. Feedback	Detailed, delayed, extrinsic
High Psychological Fidelity	
7. Persona	Distinctive, coherent
8. Emotion	Dynamic, cumulative
9. Manifestation	Variable, pervasive
Low Physical Fidelity	
10. Body	Actor photographs
11. Voice	Actor recordings
12. Input modality	Typed text

Conversely, we hypothesize that *virtual patients that emphasize functional over physical fidelity can be cost-effective for training interviewing skills.* We experimented with the decisions in Table 1. A few caveats are in order. We do not claim these as the only relevant fidelity elements. We hypothesize that they are important for training. We do not advocate these decisions for all virtual characters. For example, we made different decisions for different goals in the virtual coach that supervises role-play. We do not claim that our virtual patients uniquely manifest each decision. But they may uniquely manifest the set. We do not know how individual decisions impact training effectiveness. In fact, they are not independent. We only hypothesize that functional fidelity is critical. We do not claim that these decisions are optimal. We hypothesize a favorable cost-benefit ratio.

We applied these decisions in virtual patients in STAR™ Workshop (Hayes-Roth et al 2004, 2009), a system to train primary care clinicians in the *Engage for Change (E4C)* protocol for brief intervention in alcohol abuse (Fig 1 & 2). STAR also includes a virtual coach who provides instruction, role-play supervision, feedback, and coaching. Working in a *guided mastery* paradigm, she requires a student to interview each patient in perfect compliance with the E4C protocol before proceeding to the next one, and to do so with all 3 patients to complete the workshop.

Figure 1. STAR™ Coach Harmony and 2 Virtual Patients

A. Inform the patient of health risks:
1. Raise the topic of alcohol in a general health context.
2. Inform the patient of his/her specific health risks.

B. Acknowledge the patient's point of view:
3. Invite the patient to express concerns about health risks.
4. Accept the patient's stated concerns.

C. Encourage the patient to make a change:
5. Invite the patient to make an appropriate change step.
6. Ask the patient to commit to make the change step.

Figure 2. Engage for Change (E4C) Protocol

High Functional Fidelity

1. Free-form Natural Language User Input. During interviews, clinicians must listen to patients and generate dialogue. To simulate this requirement, our virtual patients accept free-form natural language input (Fig 3). They recognize inputs with context-sensitive pattern matching. For example, a student's question, "How do you feel about that?" matches the pattern "#ask-reaction #topic." #ask-reaction matches various phrases and #topic is instantiated by anaphoric reference. In different contexts, the patient uses this pattern to recognize inputs such as, "what are your thoughts about drinking?" or "will you try a 12-step-program? Patients' responses are also context-sensitive. For example, the patient responds differently to "will you try a 12-step-program?" based on his/her current mood, whether the student has performed E4C steps 1-4, whether a 12-step program is appropriate, etc. The patient replies (e.g., "I'll think about that" or "I don't need a lecture"), changes mood (e.g., no change or less compliant), and records dialogue and score in the database. Alternatively, low-fidelity input via multiple-choice menus is simpler and less costly. But it allows students to merely recognize or guess correct dialogue. This is much easier than listening carefully to patients and generating correct dialogue.

2. Mixed-Initiative Conversation. Clinicians converse with patients who get impatient, interrupt, change the subject, etc. To replicate these challenges, our patients do mixed-initiative conversation; student or patient can take the lead at any time. For example, in Fig 3, student Kay hesitates; patient Lee gets impatient and takes the lead. Our architecture supports this by integrating agenda-driven and interrupt-driven control of patient behavior. Alternatively, low-fidelity enforcement of strict turn taking (patient or student, but not both, initiates and the other must reply) is simpler and less costly. However, it does not train the more demanding conversation skills required on the job.

3. Adaptive Behavior. In real interviews, clinicians fly solo, without mentors giving advice, averting disasters, or getting them back on track. To simulate this challenge, we designed patients to adapt to student behavior, protocol-compliant or not, until interviews reach their natural conclusions. For example, Lee's response to the alcohol topic depends on how Kay does E4C steps 1 and 2. In Fig 2, she fails to mention the general health context and then hesitates. Feeling uncertain, Lee responds tentatively, then defiantly. In Fig 3, Kay mentions the general health context and Lee's specific health risks. Interested, Lee responds compliantly. When Kay subsequently patronizes Lee, telling him to reduce his drinking, he feels defensive and rebels. Our approach supports cumulative adaptation of patient behavior with highly contingent nested agendas that are sensitive to run-time context, including preconditions (e.g., E4C step performance) and multi-dimensional moods. Alternatively, it would be simpler to make scripted patients that channel students through ideal interviews. But that would not require students to attend closely to patients and adapt their own behavior, as they must with real patients, nor would it allow them to enjoy improvements in patient response to improvements in their protocol adherence on successive role-plays.

Lee (cheerful): I'm glad it's just a sprain. What a relief!
Kay: Can we talk a little bit about your drinking?
Lee (uncertain): Um, I guess so.
Kay: …
Lee (impatient): What did you want to talk about?
Kay: …
Lee (very impatient): Well, if that's it, I'll be leaving.

Figure 2. Kay Starts Well, Then Hesitates

Lee (neutral): Just a sprain. Well, thank goodness for that!
Kay: Can we discuss how alcohol is affecting your health?
Lee (neutral): All right.
Kay: You know, Lee, alcohol can be a factor in your ulcers.
Lee (interested): Really? That's news to me.
Kay: You need to reduce your drinking, Lee.
Lee (defensive): Whoa! You're saying I drink too much?
Kay …
Lee (impatient, angry): I don't know what you want from me.
Kay: I want to talk about your drinking.
Lee (very angry): You already said that.
Kay: Well, you're not listening.
Lee (totally angry): I didn't come for a lecture. I have to go.

Figure 2. Kay Starts Well, Then Patronizes

4. Variable Expression. Even when clinicians see continuing patients, who report similar symptoms, make similar comments, etc, they must participate mindfully in the interview, listen and observe, and generate appropriate dialogue and advice. To simulate this requirement and provide authentic practice on every role-play, our virtual patients randomly vary expression of their intentions, while remaining true to character. For example, Figs 2-3 show how Lee varies his opening line. In this case, he chooses among explicit alternatives. In other cases, patients might instantiate dialogue templates with different variable values. Alternatively, making patients play their roles the same way every time is simpler, faster for students, and less costly. But it encourages students to perform on autopilot, relying on memory for correct or incorrect dialogues, with minimal training benefits.

5. Intrinsic Reinforcement. Since the training goal is to teach a particular interview protocol, virtual patients provide intrinsic reinforcement, both positive and negative, based on a student's protocol instantiation. Figs 2-3 show Lee's graded responses to Kay based on how well she instantiates E4C protocol steps 1 and 2. This is done with the context-sensitive pattern-matching mechanism described above. Alternatively, it would be simpler to correct student errors on the fly and force interviews back on track. But that would not require students to detect and recover from their errors, as they must with real patients.

6. Detailed Delayed Feedback. The training literature shows that feedback is most effective when it is delayed, to not break the *flow* of practice, and detailed enough that students clearly see and correct errors (Druckman & Bjork 1994; Miller et al 2006). Our patients record numerical scores and student-patient dialogue, for each protocol step in each role-play. A virtual coach uses this information to give detailed post-practice feedback, as shown in Fig 4. Alternatively, injecting feedback in real time is simpler and obviates the need for a database. But it would reduce the effectiveness of both role-play and feedback.

Coach: That didn't go very well, did it, Kay? Don't worry. Most students need a few role-play sessions with their first patient. You can try again in a moment. First, let's review. You got 3 points on step 1. That's up from 1 last time. Good job, Kay. You said, "Can we discuss how alcohol is affecting your health." By raising the topic in a neutral health context, you made it easy for Lee to agree, so he said, "All right." Good performance on step 1.

Coach: Next, you said, "You know, Lee, alcohol can be a factor in your ulcers." Excellent. You told Lee how alcohol could be affecting his specific health problems. That got him interested. He said, "Really? That's news to me." Good job.
Coach: Next you said, "You need to reduce your drinking, Lee." This was good medical advice, Kay, but it was too soon to advocate change. Unfortunately, your good advice actually made Lee feel defensive and more resistant to change. So he said, 'Whoa! You're saying I drink too much?" Next time, try to move to step 3 of your protocol. Invite Lee to express his own thoughts about how alcohol is affecting his health. That will make him feel like a respected partner in the discussion and get him thinking about the possibility of making a change in his drinking.
Coach: Kay, overall, you did much better on this role-play with Lee. Nice improvement. Before you do another practice with Lee, I would like to give you a little extra coaching on step 3. OK?

Figure 4. Coach Harmony's Feedback Role-Play 2

High Psychological Fidelity

7. Distinctive Coherent Persona. Like real patients, each of our virtual patients has a distinctive, coherent persona—identity, age, gender, ethnicity, history, reason for clinic visit, red flag for alcohol abuse, rationale for drinking, readiness for change. Alternatively, it would be simpler and less costly to embed limited persona content as needed in scripted role-plays. But rounded patient personas are more engaging and support a wider range of characteristic behavior. Clinicians must adapt their interviewing to some qualities and avoid distraction by others, as they must with real patients. It also allows different patients to represent different patient demographics.

8. Dynamic Cumulative Emotion. A great challenge during intervention for alcohol abuse is that patients can respond emotionally to charged topics, clinician manner, etc. Their emotions accumulate, interact, and influence their behavior throughout an interview. Figs 2-4 illustrate how virtual patients manifest similar emotional dynamics. This is enabled by a mood mechanism that moves a patient's *current-mood* specified distances (e.g., a little/lot more/less) in a multi-dimensional space (e.g., valence, arousal, attraction). Emotions are defined as volumes in the space, e.g., interested (val>0, att>0), uncertain (val<0, att<0), impatient (val<0, att<0, aro<0), defensive (val<<0, att<<0, aro>>0). Because they share underlying dimensions, mood changes interact. For example, in Fig 3, Lee is more uncertain each time Kay hesitates, moving his current-mood further <0 on valence and attraction. Similarly, in Fig 4, Lee is interested, until Kay patronizes him, reversing his interest, moving his current-mood from >0 to <0 on valence and attraction. Alternatively, it would be easier to give patients simple transitory emotions for specific events. But this would not simulate the nuanced emotional dynamics students must manage in real patients.

9. Variable Pervasive Manifestation. Like real patients, our virtual patients express their personas and emotion in multiple ways, e.g. dialogue content, conversation style, compliance, reaction to criticism, facial expression, body language. This gives students a natural experience of virtual patients as rounded individuals.

Low Physical Fidelity

10. Body-Actor Photographs. Our virtual patients are embodied in photos of actors displaying various gestures and facial expressions. These are organized in a 2x2 matrix representing different behaviors (e.g., talk, ask, object) in different moods (e.g., interested, defensive). During interviews, patients display photos from cells representing their intended behavior type and current-mood. Since most cells contain multiple photos, patients may draw display photos at random or with probabilities reflecting a normal distribution of behaviors within cells. This same *mind-body interface* can be used to classify and choose among different animation sequences or computer graphics commands. We chose the actor-photo embodiment to contain cost, while maximizing the visual manifestation of emotion. Although more costly animation and virtual reality ostensibly offer higher physical fidelity, they do not always convey emotion (Raij et al. 2007). Video offers even higher-fidelity presentations of real human beings, but it costs more, limits interactivity, and, paradoxically, may be less effective at inducing a suspension of disbelief.

11. Voice-Actor Recordings. Similarly, our virtual patients speak in the recorded voices of actors. During interviews, they play dialogue lines drawn from a library of annotated recordings, based on intended semantic content and current-mood. Again, they may have multiple appropriate options in a given context and choose among them probabilistically. Alternatively, lower fidelity speech synthesis offers more variability of dialogue, but conveys little emotion (Dickerson 2005).

12. Typed Text Input Modality. Our patients accept typed text input, so students can construct their dialogue. Alternatively, input menus would be simpler and less costly, but that would finesse challenging requirements of real patient interviews. Spoken input would be higher fidelity and perhaps more effective, but speech recognition software would increase cost and limit access.

Training Effectiveness of STAR Workshop

A controlled experiment evaluated STAR Workshop and compared it to E-Book and Control conditions (Hayes-Roth et al 2004, 2009). E-Book students studied materials on a Web site at their own pace, including all STAR content from instructional materials and coach dialogue, plus interview transcripts illustrating correct performance and errors. Control students received no training. Subjects were 30 medical and nursing students. Written tests (below) were used to assign subjects to low, medium, and high pre-training E4C skill clusters, from which they were assigned randomly to training conditions. Thus, conditions balanced pre-training skills, as well as gender, age, ethnicity, and degree program. 11, 11, and 8 subjects completed the study for STAR, E-Book, and Control.

Subjects took written tests, pre-training, post-training, and post-delay (2 weeks). Each test had 6 unique items, 1 for each E4C protocol step, in random order. In each item, a patient interview-in-progress stopped just before the target E4C step. Fig 5 shows a sample test item for Step 1: Raise the topic of alcohol in a general health context. Two blind judges scored subjects' responses to each item, from -1 to 3: +1 for any E4C step, +1 for a correct step in context, +1 for a completely correct step, -1 for an error. Thus, total scores for each 6-item test could range from -6 to +18. Judges' ratings correlated $r = .94$, $p < .001$. Figure 6 shows scoring of different responses for step 1.

Context: Troy, 43, an attorney, is at your clinic to check his recovery from a broken collarbone suffered in a one-car crash. On his medical history, he reports that he consumes 30 drinks of alcohol per week. He also reports frequent insomnia and gastritis, which you think may be related to his alcohol consumption. Troy says: "The shoulder's much better. We're done for today, right?"
Question: What do you say next?

Figure 5. Sample Test Item for Step 1.

Before you go, Troy, I'd like to talk about how drinking might be affecting your health. (+3)
Troy, let's talk about your drinking. (+2)
How do you feel about your drinking? (+1)
Can you stay for a few more minutes? (0)
You know, Troy, you need to stop drinking. (-1)

Figure 6. Sample Scoring for Step 1 Responses

Immediately after training, STAR and E-Book subjects rated the E4C protocol, their training, their skills, and their intention to intervene, on 5-point Likert scales. Both groups gave mean ratings of 4 or 4.5 to all items, with no significant differences. STAR subjects performed better on tests than E-Book or Control, with all F-tests p < .001. STAR produced greater pre-training to post-training improvement (47% to 91%) than E-Book (49% to 61%) or Control (44% to 48%). STAR produced higher post-delay scores (89%) than E-Book (57%) or Control (45%). On all but one measure, E-Book did not differ significantly from Control. Before the post-delay written test, subjects did phone interviews with a standardized patient. Interviews were recorded and scored the same as written tests. STAR produced higher phone interview scores (92%) than E-Book (52%) or Control (51%), with all F-tests p < .001 and no significant difference between E-Book and Control.

Comments

Our study produced 2 main results. (1) It replicated the *confidence-competence gap*. Easy E-Book training inspired positive attitudes, but produced no objective improvement in skills. (2) It showed the training effectiveness of virtual patients emphasizing functional over physical fidelity. More challenging STAR Workshop training inspired positive attitudes and significantly improved performance. We obtained similar results training management communication, where STAR outperformed Menu-based virtual employees, E-Book, Control, and live workshops with human coaches and peer role-players. Together, these studies suggest that STAR Workshop may be a cost-effective approach to training communication skills. However, more research is needed to: understand the complementary benefits of virtual coach and patients, determine cost-benefit ratios for individual patient design decisions, evaluate training effectiveness for other tasks, and assess transfer of skills to on-the-job performance.

It is worth noting another risk of over-emphasizing student attitudes about training. There is "a bias that exists favoring physical fidelity over functional fidelity among the general public... physical fidelity is readily apparent and immediately gratifying. An appreciation for the analogy between how communications exist in the real world and … are preserved in [role play] is a more abstract concept that is likely to be lost on many trainees, at least initially. Increasing buy-in when functional fidelity is higher than physical fidelity will likely take time as trainees gradually experience their improvement." (Alexander et al 2005).

STAR virtual coach and patients were built with Imp™ Character Technology (Hayes-Roth 99). Besides supporting functional requirements, Imp provides a high-level authoring tool for use by domain experts, with minimal support by programmers. Dozens of researchers and students have used it to build many applications. Current projects include: Dr. Eva Hudlicka (Psychometrix) is developing a virtual coach for mindfulness meditation. Dr. Gail D'Onofrio (Yale Medical School) is developing a STAR Workshop training brief intervention in substance abuse. At Lifelike Solutions, we are developing virtual coaches for healthy eating, exercise, and other behaviors. We would be happy to discuss opportunities for other researchers to use Imp Technology in their own work.

References

Alexander AL, et al. 2005. From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in in sims & games, http://www.darwars.com.

Baer JS, et al. 2003. An Evaluation of Workshop Training in Motivational Interviewing for Addiction and Mental Health Clinicians. *Drug & Alcohol Dep*, 73:99-106.

Barney C, Shea SC. 2007. The Art of Effectively Teaching Clinical Interviewing Skills Using Role-Playing: A Primer. *Psychiatric Clinics of North America*. 30:31-50.

Barnsley L, et al. 2004. Clinical Skills in Junior Medical Officers:Comparison of Self-Reported Confidence & Observed Competence. *Med Ed*, 38:358-367.

Bracht GH, Glass GV. 1968. External Validity of Experiments. *American Educational Research Journal*.

Clark R. Building Expertise: Cognitive Methods for Training and Performance Improvement. 2003. Washington, D.C.: Int Society for Performance Imp.

Commonwealth Fund Task Force on Academic Health Centers. 2002. Training Tomorrow's Doctors: The Medical Education Mission of Academic Health Centers.

Deladisma AM, et al. 2008. Medical Student Satisfaction with Using a Virtual Patient System to Learn History-Taking and Communication Skills.

Deterding R, Milliron C, Hubal R. 2005. Virtual Pediatric Standardized Patient Application: Formative Evaluation Findings. *Studies Health Technology Inform. 111:105-107.*

Dickerson R, et al. 2005. Virtual Patients: Assessment of Synthesized vs. Recorded Speech. *E-Health: Medicine Meets VR,* 114-119.

Druckman D, Bjork RA. (1994). *Learning, Remembering, Believing: Enhancing Human Performance*. National Research Council, Washington, DC: Nat Academy Press.

Hayes-Roth B, Amano K, Saker R. and Sephton T. 2004. Training Brief Intervention with a Virtual Coach and Patients. In BK Wiederhold, G. Riva *Annual Review of CyberTherapy and Telemedicine,* 2:85-96.

Hayes-Roth B, Saker R, Amano K. 2009. Automating Brief Intervention Training with Individualized Coaching and Role-Play Practice. *Methods Med Informatics*, in rev.

Hayes-Roth B, Johnson MV, van Gent R, Wescourt K. 1999. Staffing the Web with I Characters, *Comm ACM.*

Johnsen K, et al. 2007. The Validity of a Virtual Human Experience for Interpersonal Skills Ed. *SIGCHI Conf on Human Factors in Computing Systems, 1049-1058.*

Kenny P, et al. Building Interactive Virtual Humans for Training Environments. 2007. *Interservice/Industry Training, Simulation, and Education Conference.*

Lein C, Fills CE. 2007. Using Patient-Centered Interviewing Skills to Manage Complex Patient Encounters Primary Care. *J Am Acad Nurse Practitioners*. 19:215-220.

Lok B, et al. 2006. Applying VR in Med Comm Ed: Immersive Virtual Patients. *Virtual Reality,* 10:185-195.

Lurie SJ. 2003. Raising the Passing Grade for Studies of Med Ed. *JAMA,* 290:1210-1212.

Lyles JS, Dwamena FC, Lein C, Smith RC. 2001. Evidence-Based Patient-Centered Interviewing. *J Clinical Outcomes Management*. 8:28-34.

Miller WR, Sorensen JL, Selzer JA, Brigham GS. 2006. Disseminating Evidence-Based Practice in Substance Abuse Treatment. *J Substance Abuse Treatment*, 31:25-39.

Moral RR, Salvador JJ, de Torres LP, Rados Castillejo JA. 2004. Effectiveness of a Clinical Interviewing Training Program for Fam Prac Residents. *Fam Med*, 35:489-495.

Noel GL, Herbers JE. Caplow MP. Cooper GS. 1992. How Well do Internal Medicine Faculty Members Evaluate the Clinical Skills of Residents? *Annals Int Med*, 117:757-765.

Orton E, Mulhausen P, Pangaro LN, Harvey J. 2008. Virtual Patients for Geratric Ed. *Ger & Ger Ed*. 28:73-88.

Parsons TD, et al. 2008. Objective Structured Clinical Interview Training Using Virtual Human Patient. *Med Meets VR*, 16:357.

Raij AB, et al. 2007. Comp Interperersonal Interactions with a Virtual Human to a Real Human. *IEEE Trans Visualization & Computer Graphics,* 13:443-457.

Rubel EC, Sobell LC, Miller WR. 2000. Do Cont Ed Workshops Improve Skills? Effects of a Motivational Interviewing Workshop on Substance Abuse Couns Skills & Knowledge. *Behavior Therapist,* 23, 73-77, 90.

Sanchez-Vives M, Slater. 2005. Presence -Consciousness through VR. *Nature: Neuroscience, 6:332-339.*

Scalese RJ. 2005. Features & Uses of High-Fidelity Medical Simulationss that Lead to Effective Learning: A BEME Systematic Review. *Medical Teaching*. 27:10-28.

Srinivasan M, et al. 2006. Assessment of Clinical Skills Using Sim Tech. *Academic Psychiatry*, 30:505-515.

Stevens A, et al. 2006. The use of virtual patients to teach medical students history taking and communication skills. *American J Surgery, 191, 806:811.*

Triola M, et al. 2006. A Randomized Trial Teaching Clinical Skills Using Virtual and Live Standardized Patients. *J Gen Int Med,* NY: Spinger.

Witmer B, Singer M. 1994. Measuring immersion in virtual environments, *Tech. Report 1014*, Alexandria, VA: U.S. Army Research Inst for the Beh and Social Sciences.