Published in final edited form as:

Soc Work Ment Health. 2020 ; 18(2): 121–148. doi:10.1080/15332985.2017.1336743.

Virtual Patient Simulations for Brief Assessment of Mental Health Disorders in Integrated Care Settings

Micki Washburn, PhD,

Graduate College of Social Work, University of Houston, Houston, Texas, USA

Danielle E. Parrish,

Graduate College of Social Work, University of Houston, Houston, Texas, USA

Patrick S. Bordnick, PhD, MPH

School of Social Work, Tulane University, New Orleans, Louisiana, USA

Abstract

This mixed methods pilot investigation evaluated the use of virtual patient simulations for increasing self-efficacy and diagnostic accuracy for common behavioral health concerns within an integrated care setting. A two by three factorial design was employed to evaluate three different simulated training conditions with a sample of 22 Masters level behavioral health students. Results support engagement in virtual patient simulation training to increase students' self-efficacy in brief clinical assessment, and support the use of virtual patient simulations to improve diagnostic accuracy. Results further indicate that virtual patient simulations have sufficient levels of usability and acceptability as a tool for developing brief clinical interviewing skills, and that participants found this method of instruction to be a valuable adjunct to traditional classroom or field based training. Future directions and next steps for the integration of technology enhanced simulations in clinical social services education are explored.

Keywords

virtual patients; integrated care; simulation based learning; mental health; self-efficacy

The American Psychiatric Association (2013) estimates that approximately one in four adults, and one in five children and adolescents in the United States suffer from a diagnosable mental health disorder at any given time. Overall costs associated with mental illness and substance abuse are more than those associated with diabetes, heart disease and cancer, and taken together are estimated to exceed 500 billion dollars annually (Collins et al., 2011; Jason & Ferrari, 2010). Two thirds of these costs are “indirect costs” such as lost productivity and earnings, as well as expenditures associated with involvement in other public systems, , all of which can be significantly minimized by accurate identification and appropriate treatment (Singh & Rajput, 2006).

Traditionally, American health systems have taken a segregated treatment approach to co-occurring physical health and mental/emotional health concerns (Horvitz-Lennon, Kilborne & Pincus, 2006). Primary care physicians may not routinely assess for mental health issues, and frequently, those exhibiting mental health concerns in primary care or emergency settings are referred out for mental health services. As such, when an individual exhibits both physical and mental health concerns, they are treated by different providers in different settings. Given that those living with chronic physical health conditions are at increased risk of co-occurring mental illness, particularly recurrent major depressive disorder (Reeves et al., 2011), the need to address both physical and mental health concerns concurrently is necessary to improve patient outcomes and control health care costs (Olfson, Kroenke, Wang & Blanco, 2014). However, the current lack of integration between physical and behavioral health services has led to challenges in early identification and treatment of mental illness within primary care (World Health Organization, 2008). With the Patient Protection and Affordable Care Act (2010), efforts are being made to better integrate these services in settings where many individuals initially seek help for physical complaints (Olfson, et.al, 2014). By assessing for mental health concerns in all settings in which a client may receive his or her health care, early diagnosis and service linkage can be improved.

As aspects of the Affordable Care Act are increasingly implemented, more counselors, psychologists, and in particular social workers, will be working as part of integrated primary care teams in an attempt to address physical and behavioral health concerns concurrently (Horevitz & Manoleas, 2013; Lundgren & Krull, 2014). However, the extant literature focusing on “best practices” for teaching brief mental health assessment skills within community based integrative care settings is currently lacking. Didactic classroom training focused on the recognition and evaluation of mental health disorders alone may not lead to proficiency in brief mental health assessment. Recent works across a number of allied health professions highlight the challenges of training future mental health professionals to conduct thorough, but brief, assessments in non-traditional mental health settings. These challenges result from a number of systemic factors such as: high levels of co-morbidity among mental illnesses, limited diagnostic specificity in relation to symptom presentation, and incomplete training resulting from inadequate experiential learning opportunities (Albright, Adman, Goldman & Serri, 2013; APA, 2013; McDaniel et al., 2014; O'Donnell, Williams & Kilborne, 2013; Pollard et al., 2014; Sampson, Parrish, & Washburn, 2017). In addition, classroom training alone may not build the requisite self-efficacy needed for future mental health professionals to efficiently work within fast-paced emergency and primary care settings (Holden, Meenaghan, Anastas & Metery, 2002; Horevitz & Manoleas, 2013; Pinquart, Juang & Silbereisen, 2003).

Live, supervised field based education continues to be an essential component to the development of direct practice skills in social work (Council on Social Work Education, 2015). However, there are often few practice opportunities related to brief assessment of mental health within many traditional social work field placement settings. Mental health concerns are not given primary consideration in many social work placement sites such as schools, the child welfare system, and the adult and child juvenile justice systems, even though mental health issues most certainly impact the outcomes of clients within those systems\ (Healy, Meagher & Cullin, 2007; Kelley et al., 2010; Ko et al., 2008). Although

some specialized field placements are conducted mental health specific settings such as outpatient clinics, substance abuse treatment facilities, or psychiatric hospitals, these placements infrequently include emergency or primary care settings (McDaniel et al. 2014).

Additional ethical concerns persist about having unseasoned clinicians assessing clients from vulnerable populations, regardless of the setting (Healy et al., 2009; Skovholt & Ronnestad, 2003; Theriault, Gazzola & Richardson, 2009). Social work students often enter field settings with little or no clinical training other than a few courses comprised of classroom lectures and only an introductory knowledge of basic clinical skills, such as how to conduct a thorough assessment (Wayne, Bogo & Raskin, 2010). Students are expected to learn as they go, and acquire these skill sets while actually working with clients. However the intensity and quality of training they receive as a part of this process can vary significantly from profession to profession and from setting to setting (Ellis et al., 2013). This concern becomes more acute for provisionally licensed professionals, particularly if they are practicing in settings with very high caseloads and experience inadequate clinical supervision, increasing the potential for harm to vulnerable clients (Auger, 2004; Bogo, 2015; Ellis et al., 2013; Primm et al., 2010). Thus, it is essential to explore novel ways in which instructors in social work graduate programs can further develop students' mental health assessment skills within settings in which the clients may first present, and in ways that minimize the potential for harm to live clients.

Experiential learning has long been a cornerstone of social services education (Kolb & Kolb, 2012) in both clinical social work (CSWE, 2015; Wayne et al. 2010) and applied psychology (APA, 2015). From the beginning of one's graduate program, students engage in basic experiential learning with instructor led or peer-to-peer role-plays. Lee and Fortune (2013a, 2013b) assert that educational activities using both participation and conceptual linkage have the greatest impact on student learning outcomes. Accordingly, some social work (Badger & MacNeil, 1998, 2002; Bogo, Rawlings, Katz & Logie, 2014; Rawlins, 2012;) and psychology (Levitov, Fall & Jennings, 1999; Masters, Beacham & Clement, 2015) programs have taken experiential learning beyond basic in class role plays by implementing standard patient simulations into their curriculum to assess competency.

Live practice with standardized simulated clients, usually actors, gives students the opportunity to receive structured feedback to assist in their preparation for work in the field. These simulation-based practice opportunities integrate essential components of adult learning such as being reflexive and self-directed, while engaging in training exercises that are directly relevant to one's development as a practitioner (Knowles, 1984). Live patient simulations allow students to practice key skill sets and to demonstrate their application in real world settings (Badger & McNeil, 1998, 2002; Cooper & Briggs, 2014; Petracci & Collins, 2006).

Unfortunately simulations using live actor "patients" can be costly and time consuming, as well as require a great deal of organizational coordination. Actor training is labor intensive and may need to be done repeatedly from one semester to the next, necessitating departments to incur reoccurring costs (Cook & Triola, 2009; Cook, Erwin & Triola, 2010). Coordinating the schedules of actors, instructors, and students creates additional logistical

challenges, especially when there are a large number of students engaging in these simulated learning opportunities. These challenges are amplified when trying to conduct simulations that involve students and faculty from multiple departments to accurately reproduce interactions within an integrated care team (Sampson et al., 2017). Actor based simulations can only be done one at a time and the quality of the simulation may vary from actor to actor (Triola et al., 2006). Factors such as actor fatigue affecting standardization may also come into play when actors are expected to engage in multiple simulations in a day. Furthermore, students cannot stop in the middle of the simulation and start over, or execute the simulation multiple times following self-reflection on their performance, or after feedback from instructors or peers. (Kenny, Rizzo, Parsons, Gratch & Swartout, 2007) However, incorporation of technology-based simulations, such as interactions with virtual humans, into direct practice education may be a way to minimize the drawbacks of live patient simulations (Triola et al., 2006; Washburn, Bordnick, & Rizzo, 2016).

Virtual patient simulations represent a specific type of technology enhanced simulation believed to offer many of the benefits of live actor simulations, but without the logistical drawbacks commonly associated them (Carter, Bornais, & Bilodeau, 2011; Khanna & Kendall, 2015; Triola et al., 2006). These simulations utilize virtual human agents (patients) to reproduce an interactive clinical encounter. Virtual patients are fully interactive avatars that can interact with students in many of the same ways that “actor” patients can. They can respond to student inquiries and provide information on their background, histories, and symptoms. Based on how the virtual patients are programmed, they will respond to either text input, voice input or both. In their early stages, virtual human agents were unconvincing in their appearance and movement, making them inadequate for realistic clinical simulations. However today’s virtual patients are based on innovative technology with enhanced graphics that can more accurately portray interactions with real clients (Washburn et al., 2016; Washburn & Zhou, in press). Virtual patients are customizable in terms of age, gender, race, ethnicity, and presenting issue. In addition, the practice setting in which the virtual clients are situated can also be customized, allowing for interactions with a diversity of clients and practice settings beyond what one may typically encounter in his or her field placement (Levine & Adams, 2013).

Researchers worldwide are exploring the utility of technology enhanced simulations for clinical training due to its long term cost effectiveness and ease of dissemination (Bateman, Allen, Kidd, Parsons & Davies, 2012; Cook et al. 2010). Emerging research into the exact mechanisms that make virtual patient simulations an effective training tool have found that factors such as interactivity, ease of navigation, ability to accurately depict clinical scenarios, and the availability of timely an appropriate feedback all lead to increased usability and clinical skill acquisition (Huwendiek et al., 2013; Talbot, Sagae, John & Rizzo, 2012).

Today’s virtual patients do not require a dedicated space or specialized equipment for realistic simulated encounters. They can be accessed from any personal laptop computer, and can be used multiple times per day by numerous users, reducing the time it takes for a large cohort of students to engage in simulation based learning, and eliminating the need to coordinate schedules of students, actors and faculty for live simulations. The clinical simulation hardware setup similar to the one used in the current study cost approximately

\$2,000, making it affordable for most clinical training programs. Over time, virtual patient simulations prove to be more affordable than actor based simulations as they are used year after year, and can be shared by several departments within a University or even shared between multiple Universities.

Research suggests that skills learned during virtual patient simulations will transfer over to interactions with live clients and that encounters using virtual patients result in equivalent outcomes to simulations using traditional standard actor patients (Triola et al., 2006). Moreover, students engaging in these simulations rated them as a valuable way to develop direct practice skills (Cook et al, 2010; Lichvar, Hedges, Benedict & Donihi, 2016; Washburn et al., 2016). Multiple studies have found that simulations conducted with virtual patients have comparable outcomes to those conducted with standard actor patients, and that simulation based education was superior to no training interventions (Cook & Triola, 2009; Cook et al., 2010; Kenney et al. 2007; Parsons et al., 2008).

Prior research on simulations using virtual patients demonstrated increased diagnostic accuracy for specific mental health conditions including conduct disorder, post-traumatic stress disorder (PTSD) and major depressive disorder (MDD) (Kenney et al, 2007, Parsons et al., 2008, Satter et al, 2012). Virtual patient simulations have also been associated with improvements in brief assessment skills in a sample of medical professionals a primary care setting (Satter et al., 2012).

In addition to improving identification of specific mental health conditions, simulations using virtual patients have been also been used effectively to assist physicians and psychologists with clinical history taking and the development of clinical reasoning skills (Flemming et al., 2009; Kenny et al., 2007; Pantziaras, Fors, & Ekblad, 2015; Parsons, 2015; Posel, McGee, & Fleize, 2014; Satter et al., 2012). Other studies have explored the use of virtual human technology for teaching basic practice skills (Reinsmith-Jones, Kibbe, Crayton & Campbell, 2015), building physician's empathic communication skills (Kleinsmith, Rivera-Gutierrez, Finney, Cendan & Lok, 2015) and for teaching brief intervention approaches, such as motivational interviewing, commonly used in primary care (Friedrichs, Bolman, Oenema, Gayaux & Lechner, 2014). These training simulations were highly rated by students and hold promise as effective educational interventions.

Self-Efficacy and Diagnostic Accuracy

Self-efficacy is grounded in social cognitive theory and is defined by Bandura as "people's beliefs about their capabilities to exercise control over their own level of functioning and other events in their lives" (1991, p. 257). Self-efficacy (confidence) differs from efficacy (competence) in that it refers to one's belief that one can affect an outcome versus one's actual ability to affect said outcome. The relationship between self-efficacy and skill level is often non-linear (Judge, Jackson, Shaw, Scott & Rich 2007; Larson & Daniels, 1998). Overall, prior work in this area shows that student's self-efficacy increases as a result of practice, but that self-efficacy is not always predictive of one's level of skill (Rawlings, 2012; Rawlings, Bogo & Katz, & Johnson, 2012; Kleinsmith et al., 2015). A meta-analysis conducted by Judge et al., (2007) found that for low complexity tasks, self-efficacy was

predictive of job related performance, but not for tasks that are of medium and high complexity.

As any seasoned clinician knows, a certain level of self-efficacy is necessary to engage in effective clinical work. Larson and colleagues (1999) and Rawlings and colleagues (2012) explored the impact of students' perceptions of self-efficacy in relation to the clinical skills he or she demonstrated. They found students with self-efficacy slightly higher than his or her actual skill level was optimal for experiential or simulation based learning. Conversely, students with levels of self-efficacy that were significantly higher or lower than their corresponding skill level have been found to negatively impact the acquisition of new practice skills. Larson & Daniels (1998) also found that novice clinicians with high self-efficacy experience lower levels of anxiety in relation to live clinical interactions.

Overall, self-efficacy has not been found to be significantly impacted by the type of instructional method employed. Jeffries, Woolf & Linde (2003), in a study comparing technology based instruction to traditional instruction, found there were no differences in students' perceived self-efficacy based on the type of instruction they received. Additional studies comparing technology enhanced teaching methods with traditional methods such as lectures also found no significant difference in students' self-efficacy (Alinier, Hunt, Gordon, & Harwood, 2006; McConville & Lane, 2006, Leigh, 2008). Campbell et al. (2015) noted that students who received training for clinical interviewing skills via a virtual world platform had even greater increases in self-efficacy than those receiving this training via live peer-to-peer role-play format. These findings are especially relevant in the age of online and technology enhanced education where traditional lecture based classes are becoming less common.

Author1b and colleagues (2016) conducted an initial feasibility and acceptability study of virtual patient simulations for increasing diagnostic accuracy, self-efficacy and brief interviewing skills in a sample of social work students. This study moved beyond prior research using virtual patients for mental health training, and evaluated the use of virtual patients to develop students' ability to discriminate *among* diagnoses rather than simply identify a particular diagnosis. Results supported an increase in students' diagnostic accuracy following virtual patient practice simulations. Overall, student feedback concerning the usability and acceptability of the virtual patient simulations was positive.

Given the potential promise of virtual patient simulations in social services education, the aim of this project was to further evaluate if virtual patient simulation training is an efficient and user-friendly instructional method to increase students' diagnostic accuracy and self-efficacy in brief mental health assessment. Three simulation based training approaches were assessed for within group improvement, and compared at post-test to determine if engagement in different variations of these simulations is associated with differences in diagnostic accuracy and self-efficacy.

Methods

Participants

Approval for this project was granted from the Principal Investigator's (PI) university Committee for the Protection of Human Subjects. To increase potential sample size, recruitment of graduate students in the fields of social work and psychology commenced over one semester at the PI's university. Flyers with basic study information were posted in the Social Work and Psychology departments, and graduate students in these disciplines were also emailed via departmental list serve concerning an opportunity to participate in original research which was independent of any class credit or grade. Participation criteria included: being a currently enrolled Social Work or Psychology graduate student, completion of a graduate level class in clinical assessment of mental health disorders, and being able to commit to attending up to five one-hour meetings at the College of Social Work. Potential student participants were asked to contact the principal investigator (PI) via email to set up a time for enrollment and completion of consent documentation. Of the thirty-two students who expressed an interest in the study, six declined to enroll due to the time commitment required for participation. Twenty-six students completed the consent process and four dropped out prior to randomization and initiation of protocol due to time constraints, resulting in a convenience sample of 22 students.

Design and Procedure

Participants completed the informed consent process, and were then randomized to one of three practice conditions via a computer assisted randomization program. A description of each of these practice conditions is included in the following section. Participants then completed a demographic questionnaire and participated in a 20-minute informational session on virtual patient simulations. This session included a short description of the virtual patient software, instructions how to communicate with the virtual patients, two videos showing examples of students interacting with virtual patients, and overview of the tasks they would be performing.

Overview of Study Tasks and Incentives

A randomized 2X3 mixed factorial design with repeated measures was employed as seen in Figure 1. All participants were then told that they would complete a brief (30 minute) clinical assessment interview with either a live standardized actor patient (condition 1 - SAP) or a computer based virtual patient (conditions 2 - VPP and 3 - VPN). Then, based on their condition, they would either complete a series of 3 weekly training sessions using the virtual patient simulations (conditions 1 - SAP and 2 - VPP) or receive no additional training (condition 3 - VPN). Finally, student participants would complete a post-training clinical assessment interview with the same type of patient that was used for their baseline assessment interview. All participants were given incentives in the form of a gift card following completion of each phase of the project. They received the first incentive (\$10) after completion of the initial clinical interview, the second incentive (\$20) after completion of training simulations and the third incentive (\$20) after completion of the final assessment interview. Participants who were randomized to the no training (VPN) condition received

both incentive #2 and #3 after completion of the final assessment interview. All 22 participants completed the study.

Standard Actor Patient Overview

Standard Actor Patients were graduate student actors selected from the University's drama department. They were trained by the PI and a project consultant (acting coach) to portray clients who had similar back stories and mental health concerns as the virtual patients. Four actors of different genders and ethnicities were used for the simulations, and these actors mirrored the genders and ethnicities and mental health concerns of the virtual patients that were used in the study. These mental health concerns included: clinical depression, post-traumatic stress disorder, substance abuse. Training took place at the PI's university over the course of two months where the actors would receive feedback from the consultant and the PI concerning the accuracy and consistency of their client portrayal to assure reliability of presentation across participants. The guidelines for training standard actors can be found in Bogo et al., (2014). As with the virtual patients, different standard actor patients were used for both the initial and final clinical interviews. All simulated sessions conducted with the standard actor patients were conducted in an office that resembled a traditional therapist office, where the "actor" patient would be seated on the couch facing the participant.

Virtual Patient Overview

The virtual patients in this study were developed by the University of Southern California Institute for Creative Technologies and used with their permission. This software runs on a standard desktop or personal laptop computer with enhanced graphics capabilities. The software package used in this study featured six virtual patients of various ages, genders, and ethnicities to vary the clinical experience. Each virtual patient had a specific set of symptoms and life experiences related to a common mental health disorder such as anxiety, depression, post-traumatic stress or substance abuse. They appeared in a setting that resembled a traditional therapist's office and sat on a couch facing the participant. The virtual patients were projected on to a screen to make them appear to be life size and thus, seem more realistic. Figure 2 shows one of the virtual patients used for this study.

Communication with the virtual patients was done via text based interfacing. Participants would speak to the virtual patient while the PI would simultaneously type participants' verbal responses into the VP program via laptop, thus prompting a verbal response from the virtual patient. A text box of the virtual patient's verbal response would also be generated to keep a running log of the interaction between the participant and the virtual patient that could be downloaded for later review.

Baseline Clinical Assessment Interviews

Participants were told that they were working as part of an integrated care team within a local public health system. The "patient" had just seen his/her primary care physician for a routine physical and the physician had requested the social services provider to conduct a brief mental health assessment to screen for any co-occurring behavioral health conditions. Since this screening occurred in a busy primary care setting, participants would have a maximum of 30 minutes in which to conduct the assessment. Participants were informed

that they could take notes during the clinical interview and were given a one page assessment guide to more closely approximate a real clinical encounter in a community based setting. Similar guidelines are commonly used in clinical practice to assess for mental health disorders (APA, 2013; First, Williams, Karg & Spitzer, 2015).

Following the interview, participants completed a Diagnostic Rating Form (DRF) where they would any applicable DSM diagnosis for the patient. A copy of the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (APA, 2013) was available to reference while completing the DRF. Participants were informed that more than one diagnosis may be applicable and that potential diagnoses were limited to the following categories: Anxiety Disorders, Substance Use Disorders, Trauma and Stress Related Disorders, Depressive Disorders and Bipolar Disorders. Finally, participants completed a measure of self-efficacy related to the tasks of clinical interviewing and brief diagnostic assessment. A more detailed description of the Diagnostic Rating Form and the Social Work Self-Efficacy Measure are included in the Measures section.

Virtual Patient Simulation Practice Sessions

Following the initial assessment, students engaged in either three sessions of virtual patient simulation practice or no additional simulation practice, depending on the training condition to which they were assigned. Participants assigned to the Standard Actor with Practice group (SAP), or the Virtual Patient with Practice group (VPP) engaged in three consecutive weekly 45-minute training sessions. Prior to each training session, participants were again given information on how to interact with the virtual patients, and a copy of the same clinical interview guide they had used for their baseline interviews. During these training sessions, participants would engage in a simulated clinical assessment interview with a virtual patient to practice their brief assessment skills. At the end of each practice session, they were asked to render a DSM diagnosis for the virtual patient they had just interacted with. During training, participants received concurrent and terminal feedback from a mental health professional with over 10 years of direct practice experience concerning the quality of their interactions with the virtual patients, and potential DSM diagnoses for each.

Participants had the opportunity to interact with a total of four different virtual patients (one during the first session, one in the second session and two in the final session) though out the course of the training protocol. The order of presentation and patients used for the practice simulations were randomized to control for differences in the quality of virtual patients used for this study. Participants in the Virtual Patient with no Practice group (VPN) condition received no additional simulation training after the initial assessment interview.

Post Training Clinical Assessment Interviews

Approximately one month after the baseline assessment interview, all participants then completed a second brief clinical assessment interview with another standardized actor patient or virtual patient. Participants again completed the DRF, a self-efficacy assessment, as well as an anonymous online usability survey (UFF), which included an additional three open-ended questions concerning their experience with the virtual patient simulations. The order and character used for the patient interviews was randomized for each participant to

control for any difference among virtual patients or actor patients, and novel patients were used for both the baseline and final clinical interviews.

Assessments & Measures

Demographic Information (DI): Participants were asked to complete a standard demographic questionnaire. Demographic information collected included age, gender, race/ethnicity, type of graduate program (psychology or social work), and years in program.

Diagnostic Reporting Form (DRF).—The DRF is based on DSM-5 criteria (APA, 2013). Students' DRF assessments were scored on a scale of 0 (low diagnostic accuracy/inadequate information gathered) to 50 (high diagnostic accuracy/necessary information was gathered). Participants were scored on the following assessment items: 1) History taking - obtain specific client symptom history and relevant biopsychosocial assessment information (20 points); 2) Diagnosis - provide a DSM diagnosis as appropriate (10 points); and 3) Case conceptualization - provide a brief justification/summary of the diagnosis/es and any additional diagnoses that are considered but ultimately ruled out (20 points). To determine if the virtual patient simulation training has an impact on students' diagnostic accuracy specifically, scores on the second section of the DRF concerning "diagnostic accuracy" were then converted to ordinal variables based on the following criteria: completely incorrect (0) – participant did not identify any applicable diagnosis/es; partially correct (1) – participant correctly identified at least one of the applicable diagnoses; or completely correct (2) – the participant correctly identified all of applicable diagnoses. Prior research studies on clinician learning have supported the validity of using similar rating/scoring methods assessing learning on clinical diagnostic tasks using standard patients (Feingold, Calalupe, & Kallen, 2004; Masters et al., 2015). A PhD level clinician independent of the university with over 10 years of direct mental health practice and supervision experience received training on DRF scoring, and served as the independent rater.

Social Work Skills Self Efficacy (SWSSE)—The social work skills self-efficacy measure is a self-report measure through which students' rate their own levels of proficiency in relation to a variety of tasks performed during the clinical interview. The current 10 item version was constructed based on Bandura's (2006) guidelines for measuring self-efficacy, and used questions from the original 52 item multidimensional instrument for social work proposed by Holden et al., (2002). Cronbach's alpha for the SWSSE total scale and subscales ranged from .86 to .97. Respondents were asked to indicate her/his level of confidence in her/his current ability to perform 10 tasks specifically related to clinical interviewing in a manner that an experienced supervisor would evaluate as excellent. Ratings ranged from 0 = I am not at all confident I can execute this skill, 5 = I am moderately confident I can execute this skill, and 10 = I am very confident that I can execute this skill. Scores on each question are summed to determine the total SWSSE score with a minimum total score of 0 and a maximum total score of 100. This measure is included in the Appendix.

Usability Feedback Form (UFF).—A standardized evaluation form assessed student preferences and perceptions of usability concerning the virtual patient simulations. The form is a modified version of the System Usability Scale (Brook, 1996) and explored student

experiences with the VP method. This scale was validated for use with multiple forms of technology-based interactions, and has an average reliability of $\alpha=.85$ (Bangor, Kortum & Miller, 2008). Students rate their experiences on a scale of 1 (low usability) to 5 (high usability) for ten questions related to the overall value of method, and the perceived effectiveness of method in preparing participants for actual clinical interviews. Directions for scoring and evaluation are presented in Bangor et al. (2008). Scale scores above 70 are considered “good”. Participants were also asked supplemental questions describing their experiences with the virtual patients, and how similar or dissimilar they were to actual clients seen in their field placements, and if they would recommend this type of training to other students.

Planned Analyses

Comparisons of pre- and post-training DRF and SWSSE scores were then computed using paired t-tests to determine if practice simulations resulted in increased overall diagnostic skills and self-efficacy on brief clinical interviewing for each group. Between groups comparisons were conducted via a one way analysis of variance (ANOVA) for the mean change scores on the DRF and SWSSE at post-test. Frequency and percentages of correct diagnoses for each group at baseline and post training were computed. Usability scores were calculated based on the total score of the UFF. Finally, Pearson correlations for changes in diagnostic accuracy, self-efficacy and usability were calculated to determine the strength of association among these variables. Open-ended feedback regarding the usability of virtual patients was analyzed using summative content analysis techniques. . Key words were identified by two coders who were independent of the project and had qualitative data analysis experience. Counts were tabulated to highlight which responses occurred most frequently.

Results

Demographics

Data were cleaned and checked for violation of assumption of normality of dependent measures As shown in Table 1, this sample had a mean age of approximately 29 year old ($SD = 6.71$ years), was ethnically diverse, predominately female, with most participants having completed two years of graduate training. Because some of the students in this sample were enrolled in the part time or weekend programs which generally take three years to complete, a mean of greater than two years of graduate education was reported. . To evaluate if there were any significant differences between groups on demographic characteristics at baseline, initial testing was done using a series of chi-square analyses for nominal demographic variables. Ethnicity variables were dichotomized into White and non-White. No significant differences in demographic characteristics were found among participants at baseline in the three simulation conditions. One way ANOVAs were used to assess for any baseline differences between groups for scale variables. No significant difference between groups were found for self-efficacy. Baseline scores on diagnostic accuracy for those completing their baseline assessment interviews with a virtual patient tended to be lower than those whose baseline assessment interview was done with a live

actor patient. As such, change scores as well as post-training scores were used to evaluate changes in diagnostic accuracy.

Diagnostic Accuracy

Paired samples t-test was performed for each simulation group's DRF scores at baseline and post-training to determine if participation in training simulations was associated with higher post training diagnostic accuracy. The t-test is robust in relation to assumptions of normality, even when sample sizes are small (de Winter, 2013). Effect sizes were then calculated using Cohen's (1992) *d* using procedure suggested by Morris and DeShon (2002) to correct for dependence between means, then adjusted for small sample size using the calculation for Hedge's *g* (Turner & Bernard, 2006).

$$d_{adj} = d \left[1 - \left\{ 3 / (4n - 9) \right\} \right].$$

As seen in Table 2, these results indicate that there was a significant increase in pre/post diagnostic accuracy for the participants in the VPP group, but not for those in the SAP group, nor the VPN group. The magnitude of the pre/post change in the VPP group was large based on Cohen's (1992) standards. To determine if there were any differences in overall change in pre/post diagnostic accuracy among simulation groups a one way Analysis of Variance (ANOVA) was conducted. No difference in DRF change scores were found among the three groups, $F(2,19) = 1.35, p = .28$.

As seen in Table 3, students who participated in the virtual patient simulation training (SAP and VPP) were able to partially or fully identify the correct diagnosis following the practice simulations. However, those who did not receive simulation training (VPN) did not exhibit the same gains in diagnostic accuracy, still rendering a completely incorrect diagnosis 25% of the time at post. Participants who conducted their assessment interviews with standard actor patients had higher percentages of correct diagnoses at both pre and post training than those who conducted their interviews with virtual patients.

Self-Efficacy

Cronbach's alpha for the ten item SWSSE used in this study was in the "good range" at $\alpha .85$ (Kline, 2000). As seen in Table 4, pre/post interviewing self-efficacy increased significantly for both the virtual patient with practice (VPP) and standard actor patient with practice (SAP) groups. Both groups exhibited large increases in self-efficacy scores as described by Cohen (1992), with the VPP group reporting the largest magnitude increase in self-efficacy. No significant difference was found between baseline and post training self-efficacy scores for students in the no training condition. There were no significant difference found among the three groups for change in self-efficacy scores, $F(2, 19) = 1.64, p = .22$.

Finally, a series of Pearson correlations were conducted to determine if there was any significant relationship between performance (DRF scores) and self-efficacy. There was no significant correlation between baseline DRF and self-efficacy scores, $r = -.32, p = .14$ nor between post-training DRF and self-efficacy scores, $r = .10, p = .65$.

Usability

Usability of the virtual patient simulation software was evaluated by calculating the overall scale score for the Usability Feedback Form (UFF). The 10 item UFF exhibited very good internal consistency reliability ($\alpha = .88$). As shown in Table 5, the overall usability of the virtual patient software was in the “good” range with a mean usability score of $M = 70.23$, $SD = 12.60$. Scores ranged from 50 to 97.50.

Based on the results of a one-way ANOVA, no significant differences were found among the three groups in relation to the overall usability score, $F(2, 19) = .45$, $p = .64$. Additionally when independent sample t -tests were performed, no significant difference in overall usability were found based on participant gender, $t(20) = -1.03$, $p = .32$, ethnicity, $t(20) = -.25$, $p = .81$, or discipline, $t(20) = -.77$, $p = .45$. When groups were collapsed to test if there were differences in usability based on whether one completed additional virtual patient training simulations, results show that those who received additional training did not have usability scores that were significantly higher than those who did receive training, $t(20) = -.85$, $p = .41$. Finally, a series of Pearson correlations were conducted to determine if usability was significantly correlated to diagnostic accuracy or self-efficacy post training. Results indicated that usability scores were not significantly correlated with post training DRF scores, $r = -.13$, $p = .58$ or with post training self-efficacy scores $r = .08$, $p = .96$, indicating that usability ratings were independent of task performance or self-efficacy.

Fourteen participants (64%) offered responses to the optional open-ended questions concerning their interactions with virtual clients, and all participants had completed at least two out of a possible five virtual patient sessions. . When participants were asked how their interactions with the virtual patients were similar or different to those with live patients in their field placements, the majority of them ($n = 11$, 78.6%) reported that these particular virtual clients were more challenging to elicit information from than live clients, as they did not spontaneously offer up information as many “actual” clients did. Respondents ($n = 9$, 64.3%) also stated that these virtual patients were very similar to real patients in terms of symptom presentation; however they were different in the fact that they did not respond to empathic statements from the interviewer like actual clients, putting the interviewer ill at ease. Finally respondents ($n = 10$, 71.4%) stated that the breadth of topics that the virtual patients could discuss was more limited than the breadth of topics that could be discussed with an actual live client. When asked how the virtual patient software could be improved, all respondents ($n = 14$, 100%) stated that inclusion of a voice recognition component, in place of manual the text input, would increase the authenticity of the simulations, as well as their emersion in them. Similarly, the majority of respondents ($n = 12$, 85.7%) indicated that inconsistencies in some of the virtual patients’ responses had a negative impact on the success of the simulations. Finally when asked about which specific tasks that the interaction with the virtual patients helped them improve upon, all those who engaged in the practice sessions ($n = 13$, 92.8%) stated that this type of simulation helped them to learn how to conduct the assessment within the 30 minute time period, and many ($n = 12$, 85.7%) also stated that they found the simulations to be helpful for learning how to inquire about the patient’s mental health symptoms.

Discussion

Overall, these results indicate that simulation based training with virtual patients may be a promising approach for increasing students' self-efficacy, and diagnostic accuracy during brief clinical assessment interviews. Results also indicate that the virtual patients used in this study had adequate levels of usability independent of users' age, ethnicity or gender, and that a students deemed these technology enhanced simulations to be helpful in developing real world practice skills. Although preliminary, the results in this study are similar to prior studies comparing outcomes of brief interviews using virtual patients to those of interviews using standard actor patients (Cook et al. 2010; Triola et al., 2006). When each training group is considered individually, the largest gain in overall assessment skills as measured by DRF scores was found in the group that used the virtual patients for both the clinical assessment interviews and additional training sessions. Both groups who completed their baseline clinical assessment interviews with virtual patients tended to have lower baseline DRF scores, indicating that it completing assessment interviews with virtual patients may initially be more difficult than completing them with live patients. However, by posttest, the scores for the group completing their interviews with virtual patients who also completed additional training experiences with the VPs (VPP) had post-test scores that were similar to those of students who completed their assessment interviews with live actor patients (SAP). However, the group completing assessment interviews using virtual patients but did not have additional practice/training opportunities (VPN) did not show the same improvements as measured by effect size. Thus, it appears that the more one interacts with the virtual patients, the larger the potential gains in diagnostic accuracy. This is a key point to consider as one of the main advantages of virtual patient simulations is the fact that these simulations can be run repeatedly, as any time of day or night, at the convenience of the student, allowing for maximal use.

Students who completed the virtual patient simulation training had on average a 13-point gain in self-efficacy, indicating that their self-efficacy increased from the moderate to high range, whereas the self-efficacy for students who did not complete the additional training simulations stayed in the moderate range. Effect sizes were large for both of these groups. Given that self-efficacy is an important component of competent clinical work, these results are promising in terms of the potential impacts that virtual patient simulations may have in relation to improving students' confidence around conducting brief clinical assessment interviews. The finding that self-efficacy was not associated with increased diagnostic accuracy was consistent with prior literature investigating the relationship between student competence and student self-efficacy (Rawlings et al., 2012; Leigh, 2008; Larson & Daniels, 1998). Although self-efficacy is an important part of competence, self-efficacy does not appear to predict competence, nor does competency increase proportionally with increases in self-efficacy. Self-efficacy appears to improve more rapidly than competence, thus educators should be cautioned to help students to accurately and objectively evaluate their current skill levels, as to not overestimate their abilities to perform various behavioral health tasks.

The usability of the virtual patients was rated as "good" and usability ratings were independent of diagnostic accuracy and self-efficacy. Similarly, differences in usability rating were not related to specific demographic characteristics of the participants. However,

the large range of overall usability scores in this sample is indicative of additional factors that may not have been captured by the UFF impacting students' perceptions of usability. It is possible that some participants were more comfortable with technology in general, which carried over to their evaluations of the virtual patient software as a clinical training tool.

As a preliminary investigation, the current study utilized a small sample of student volunteers, and a number of students who indicated interest in the study ultimately did not participate due to the out of class time commitment. Future projects may be enhanced by using virtual patient simulations as part of normal classroom instruction on clinical interviewing and the identification of mental health issues. Increasing the number of students from which data could be drawn would allow for more robust comparisons between conditions, and would serve to improve generalizability of results to social service graduate students as a whole. Similarly, it is currently unclear if three brief simulation-training sessions were enough to produce sustained changes in diagnostic accuracy and brief clinical interviewing skills. Future investigations may seek to vary the length and frequency of the practice simulations to assess the optimal "dose" of training necessary to achieve sustained skill improvement. Although the percentages of students who could correctly identify the client's mental health diagnosis was significantly higher for those completing practice simulations than those who did not, these results are limited given the sample size.

The use of text, rather than voice input, is another potential limitation of this study. Even though text input led to higher levels of accuracy than the current voice input software could attain (Washburn et al., 2016), it is possible that this method of communication may have negatively impacted students' performance on key outcome measures and somewhat limited the authenticity of the clinical encounters. It is highly recommended that future investigations utilizing virtual patients include a highly accurate voice recognition component for communication with the virtual patients, along with further refinement of the virtual patients themselves in terms of emotional responsiveness and breadth of backstory. As such, the design of future virtual patients should not simply focus on the process of relaying information, but also on the interpersonal aspects of the clinical interaction within the simulated session.

It appears that the adapted self-efficacy scale used in this study (SWSSE) demonstrated adequate internal consistency reliability as well as face and content validity. However this adapted measure has not yet been validated with a large sample. Similarly, although the DRF appeared to adequately capture important aspects of differential diagnosis based on the DSM-5 systems, further investigation into these measures is warranted to ensure validity and reliability.

Although investigations into the use of virtual patients in professional social services education is in its nascent stages, and the small sample size limits our ability to draw definitive conclusions about the use of virtual patient simulations to improve diagnostic accuracy and brief interviewing skills, these data extend our current understanding about the general feasibility and potential uses of virtual patients within social services education.

These types of simulation based learning activities have great potential for use in hybrid or fully online clinical social work training programs. Virtual patient simulations could provide students additional opportunities for client interactions which are not limited by time of day or availability of live clients and appropriate supervision. These simulations could potentially include opportunities for students to assess for specific clinical issues such as intimate partner violence or child abuse and neglect, in addition to training them in the recognition of common mental health disorders. Since the age, gender, race and ethnicity of virtual patients is customizable, virtual patient simulations could further be used to build competence with vulnerable client populations such as recent immigrants, clients engaged in the coming out process older adults needing long term care.

There is also great potential for virtual patients to be used as a continuing education tool for clinical social workers, especially when they begin a new jobs providing services for a population that they have little prior experience with. These simulations could assist in building their self efficacy in relation to a new client population and also assist them with developing context specific skill sets for work in health care settings.

In addition to their utility for individual simulations, there is potential for using of virtual patients for group based simulations. These simulations would be particularly helpful for training students from different disciplines to work as part of integrated care teams. Group based virtual patient simulations could assist with the development of key skills required in integrated care settings such as shared decision making and interprofessional communication. Groups of students could engage in these simulations simultaneously, stop to discuss options at each step of the simulation, and then make a joint decision as how to best proceed. These group simulations could also serve as opportunities for peer based feedback on group performance.

This project was innovative in a number of ways that bear mentioning. It was the first to utilize virtual patients to train future behavioral health professionals to discriminate *among* multiple potential diagnoses rather than using them to train students to recognize a *particular* diagnosis. This is an important distinction, given the co-occurrence of multiple mental health issues for many clients and the significant overlap in symptoms across diagnostic categories. Second, other than the initial pilot study conducted in relation to the current project, this is the first project to virtual patients simulations with a sample that included social workers, extending our knowledge about the use of virtual patient simulations with students in nursing, psychology and medicine. Finally, this project was innovative in that it focused on *brief* mental health assessment in an integrated care setting. Most psychosocial or initial diagnostic interviews take between 60 and 90 minutes to conduct, which is not feasible in many emergency and primary care settings in which people first present with behavioral health concerns. Thus, focusing on increasing both accuracy and efficiency in brief clinical interviewing is essential to skillful assessment and ultimately connecting clients to mental health specific services quickly and easily within integrated systems of care. Future virtual client simulations could be enhanced through the integration of 3-D technology to increase the realism of the clinical encounter. With the increased accessibility and affordability of virtual reality based technology through tools such as the Oculus Rift or the HTC Vive, educators have even more tools available to increase the realism of clinical encounters with

virtual patients. One should also consider the development of virtual patient simulations for use with mobile based platforms such as a smart phone or tablet paired with Samsung Gear VR, which could serve as a way to give students easy access to fully immersive simulations, but at a lower cost. Similarly, creation of simulations within open source platforms such as *Second Life* would provide another alternative for programs that may have limited access to the necessary software, or those with limited financial resources (Campbell et al., 2015; Levine & Adams, 2013).

Now is the time for professional accreditation bodies such as the American Psychological Association (APA), the Council on Social Work Education (CSWE) and the Council for the Accreditation of Counseling and Related Educational Programs (CACREP) to consider the development of a virtual patient repository which could be shared among multiple programs, thus decreasing cost, increasing access to a greater diversity of virtual patients, and supporting interprofessional education. Most importantly, all of this could be done without the potential of harm to actual clients in need of services. Integrating technology into the education of tomorrow's social work health care work force has the potential of enhancing learning outcomes and most importantly, the potential for improving client care.

Acknowledgments

Funding

This work was supported in part by NIH/NIDA, 1 R24 DA019798.

Appendix Social Work Skills Self-Efficacy Rating

Please use the following measure to rate how confident you are that you can do the following in a way that an experienced supervisor would rate as excellent

1. Initiate and sustain empathic, culturally sensitive, nonjudgmental, disciplined relationships with clients?
2. Elicit and utilize knowledge about historical, cognitive, behavioral, affective, interpersonal, and socioeconomic data and the range of factors impacting upon client to develop biopsychosocial assessments and plans for intervention?
3. Conduct the assessment in the time given?
4. Understand the dialectic of internal conflict and social forces in a particular case?
5. Maintain self-awareness in practice, recognizing your own personal values and biases, and preventing or resolving their intrusion into practice?
6. Practice in accordance with the ethics and values of the profession?
7. Reflect thoughts and feeling to help clients feel understood?
8. Inquire about symptoms client is experiencing in relation to his or her mental or emotional well-being?

9. Define the client's problems in specific terms?
  10. Render a correct DSM-5 diagnosis for the client?
- 0 = I am not at all confident that I can execute this skill
- 5 = I am moderately confident that I can execute this skill
- 10 = I am very confident that I can execute this skill

References

- Albright G, Adam C, Goldman R, & Serri D (2013). A game-based simulation utilizing virtual humans to train physicians to screen and manage the care of patients with mental health disorders. *Games for Health Journal*, 2, 269–273. [PubMed: 26196927]
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Alinier G, Hunt B, Gordon R, & Harwood C (2006). Effectiveness of intermediate-fidelity simulation training technology in undergraduate nursing education. *Journal of Advanced Nursing*, 54, 359–369. [PubMed: 16629920]
- Auger RW (2004). What we don't know CAN hurt us: Mental health counselor's implicit assumptions about human nature. *Journal of Mental Health Counseling*, 26, 13–24.
- Author 1b. (2016). A pilot feasibility study of virtual patient simulation to enhance social work students' brief mental health assessment skills. *Social Work in Health Care*, 55(9), 675–693. doi:10.1080/00981389.2016.1210715 [PubMed: 27552646]
- Badger L, & MacNeil G (2002). Standardized clients in the classroom: A novel instructional technique for social work educators. *Research on Social Work Practice*, 12, 364–374.
- Badger LW, & McNeil G (1998). Rationale for utilizing standardized clients in the training and evaluation of social work students. *Journal of Teaching in Social Work*, 16, 203–218
- Bandura A (1991). Social cognitive theory of self-regulation. *Organizational behavior and human decision processes*, 50, 248–287.
- Bandura A (2006). Guide for constructing self-efficacy scales . In Urdan T & Pajares F (Eds.) *Self-Efficacy Beliefs of Adolescents* (pp. 307–337). Greenwich, CT: Information Age.
- Bangor A, Kortum PT, & Miller JT (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24, 574–594.
- Barnett SG, Gallimore CE, Pitterle M, & Morrill J (2016). Impact of a paper vs virtual simulated patient case on student-perceived confidence and engagement. *American journal of pharmaceutical education*, 80(1), 16. [PubMed: 26941442]
- Bogo M (2015). Field education for clinical social work practice: Best practices and contemporary challenges. *Clinical Social Work Journal*, 43, 317–324.
- Bogo M & Rawlins M, Katz E & Logie C (2014). Using simulation in assessment and teaching: OSCE adapted for Social Work. Council on Social Work Education: Author.
- Brooke J (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189 (194), 4–7.
- Campbell AJ, Amon KL, Nguyen M, Cumming S, Selby H, Lincoln M, ... & Goncz A (2015). Virtual world interview skills training for students studying health professions. *Journal of Technology in Human Services*, 33, 156–171.
- Carter I, Bornais J, & Bilodeau D (2011). Considering the use of standardized clients in professional social work education. *Collected Essays on Learning and Teaching*, 4, 95–102.
- Cohen J (1992). A power primer. *Psychological Bulletin*, 112, 155–159. [PubMed: 19565683]
- Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS, . . . Fairburn C (2011). Grand challenges in global mental health. *Nature*, 475, 27–30. [PubMed: 21734685]

- Cook D, Erwin P, & Triola M (2010). Computerized virtual patients in health professions education: A systematic review and meta-analysis. *Academic Medicine*, 85, 1589–1602. [PubMed: 20703150]
- Cook D, & Triola M (2009). Virtual patients: a critical literature review and proposed next steps. *Medical Education*, 43(4), 303–311. [PubMed: 19335571]
- Cooper LL, & Briggs L (2014). Transformative learning: Simulations in social work education. Retrieved from <http://research-pubs@uow.edu.au>
- Council on Social Work Education (2015). Educational policy and accreditation standards. Washington, DC: Author.
- de Winter JC (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation*, 18, 1–12.
- Dudding CC, Hulton L, & Stewart AL (2016). Simulated Patients, Real IPE Lessons When you bring together students from three disciplines to treat virtual patients, interprofessional learning gets real. *The ASHA Leader*, 21(11), 52–59
- Ellis MV, Berger L, Hanus AE, Ayala EE, Swords BA, & Siembor M (2013). Inadequate and harmful clinical supervision: Testing a revised framework and assessing occurrence. *The Counseling Psychologist*, 42, 434–472.
- Feingold CE, Calaluce M, & Kallen MA (2004). Computerized patient model and simulated clinical experiences: Evaluation with baccalaureate nursing students. *Journal of Nursing Education*, 43, 156–163. [PubMed: 15098909]
- First MB, Williams JB, Karg RS & Spitzer RL (2015). User's guide to structured clinical interview for DSM-5 disorders (SCID-5-CV) clinical version. Arlington, VA: American Psychiatric Publishing.
- Fleming M, Olsen D, Stathes H, Boteler L, Grossberg P, Pfeifer J, . . . Skochelak S (2009). Virtual reality skills training for health care professionals in alcohol screening and brief intervention. *The Journal of the American Board of Family Medicine*, 22, 387–398. [PubMed: 19587253]
- Forsberg E, Ziegert K, Hult H, & Fors U (2016). Assessing progression of clinical reasoning through virtual patients: An exploratory study. *Nurse education in practice*, 16(1), 97–103. [PubMed: 26482401]
- Friederichs S, Bolman C, Oenema A, Guyaux J, & Lechner L (2014). Motivational interviewing in a web-based physical activity intervention with an avatar: Randomized controlled trial. *Journal of Medical Internet Research*, 16(2):e48, doi: 10.2196/jmir.2974 [PubMed: 24550153]
- Healy K, Meagher G, & Cullin J (2009). Retaining novices to become expert child protection practitioners: Creating career pathways in direct practice. *British Journal of Social Work*, 39, 299–317.
- Heinrichs L, Dev P, & Davies D (2017). Virtual environments and virtual patients in healthcare. *Healthcare Simulation Education: Evidence, Theory and Practice*.
- Holden G, Meenaghan T, Anastas J, & Metrey G (2002). Outcomes of social work education: The case for social work self-efficacy. *Journal of Social Work Education*, 38, 115–133.
- Horevitz E, & Manoleas P (2013). Professional competencies and training needs of professional social workers in integrated behavioral health in primary care. *Social Work in Health Care*, 52, 752–787. [PubMed: 24028739]
- Horvitz-Lennon M, Kilbourne AM, & Pincus HA (2006). From silos to bridges: meeting the general health care needs of adults with severe mental illnesses. *Health affairs*, 25(3), 659–669 [PubMed: 16684729]
- Huwendiek S, Duncker C, Reichert F, De Leng BA, Dolmans D, van der Vleuten CP, . . . Tönshoff B (2013). Learner preferences regarding integrating, sequencing and aligning virtual patients with other activities in the undergraduate medical curriculum: A focus group study. *Medical Teacher*, 35, 920–929. [PubMed: 24003850]
- Jason LA, & Ferrari JR (2010). Oxford House recovery homes: Characteristics and effectiveness. *Psychological Services*, 7, 92–102. [PubMed: 20577571]
- Jeffries PR, Woolf S, & Linde B (2003). Technology-based vs. traditional instruction: A comparison of two methods for teaching the skill of performing a 12-lead ECG. *Nursing Education Perspectives*, 24, 70–74. [PubMed: 12743975]

- Judge TA, Jackson CL, Shaw JC, Scott BA, & Rich BL (2007). Self-efficacy and work-related performance: the integral role of individual differences. *Journal of Applied Psychology*, 92, 107–127. [PubMed: 17227155]
- Kelly MS, Berzin SC, Frey A, Alvarez M, Shaffer G, & O'Brien K (2010). The state of school social work: Findings from the national school social work survey. *School Mental Health*, 2, 132–141
- Kenny P, Rizzo AA, Parsons TD, Gratch J, & Swartout W (2007). A virtual human agent for training novice therapist clinical interviewing skills. *Annual Review of Cyber Therapy and Telemedicine*, 5, 77–83.
- Khanna MS, & Kendall PC (2015). Bringing technology to training: Web-based therapist training to promote the development of competent cognitive-behavioral therapists. *Cognitive and Behavioral Practice*, 22, 291–301.
- Kleinsmith A, Rivera-Gutierrez D, Finney G, Cendan J, & Lok B (2015). Understanding empathy training with virtual patients. *Computers in human behavior*, 52, 151–158. [PubMed: 26166942]
- Ko SJ, Ford JD, Kassam-Adams N, Berkowitz SJ, Wilson C, Wong M, ... & Layne CM (2008). Creating trauma-informed systems: child welfare, education, first responders, health care, juvenile justice. *Professional Psychology: Research and Practice*, 39, 369–404.
- Kolb AY, & Kolb DA (2012). Experiential learning theory In Steel RM (Ed.) *Encyclopedia of the Sciences of Learning* (pp. 1215–1219). New York, NY: Springer.
- Knowles M (1984) *Andragogy in action: Applying modern principles of adult education*. San Francisco, CA: Jossey-Bass.
- Larson LM, Clark MP, Wesely LH, Koraleski SF, Daniels JA, & Smith PL (1999). Videos versus role plays to increase counseling self-efficacy in pre-practica trainees. *Counselor Education & Supervision*, 38, 237–248.
- Larson LM, & Daniels JA (1998). Review of the counseling self-efficacy literature. *The Counseling Psychologist*, 26, 179–218.
- Lee M, & Fortune AE (2013a). Do we need more “Doing” activities or “Thinking” activities in the field practicum? *Journal of Social Work Education*, 49, 646–660.
- Lee M, & Fortune AE (2013b). Patterns of field learning activities and their relation to learning outcome. *Journal of Social Work Education*, 49, 420–438.
- Leigh GT (2008). High-fidelity patient simulation and nursing students' self-efficacy: a review of the literature. *International Journal of Nursing Education Scholarship*, 5, 1–17.
- Levine J, & Adams RH (2013). Introducing case management to students in a virtual world: An exploratory study. *Journal of Teaching in Social Work*, 33, 552–565.
- Levitov JE, Fall KA, & Jennings MC (1999). Counselor clinical training with client-actors. *Counselor Education & Supervision*, 38, 249–263.
- Lichvar AB, Hedges A, Benedict NJ, & Donihi AC (2016). Combination of a Flipped Classroom Format and a Virtual Patient Case to Enhance Active Learning in a Required Therapeutics Course. *American Journal of Pharmaceutical Education*, 80(10), 175. [PubMed: 28179724]
- Lundgren L, & Krull I (2014). The Affordable Care Act: New opportunities for social work to take leadership in behavioral health and addiction treatment. *Journal of the Society for Social Work and Research*, 5, 415–438.
- Masters KS, Beacham AO, & Clement LR (2015). Use of a standardized patient protocol to assess clinical competency: The University of Colorado Denver comprehensive clinical competency examination. *Training and Education in Professional Psychology*, 9, 170–174.
- Maicher K, Danforth D, Price A, Zimmerman L, Wilcox B, Liston B, . . . Way D (2016). Developing a Conversational Virtual Standardized Patient to Enable Students to Practice History-Taking Skills (1559–2332).
- McConville SA, & Lane AM (2006). Using on-line video clips to enhance self-efficacy toward dealing with difficult situations among nursing students. *Nurse Education Today*, 26, 200–208. [PubMed: 16300862]
- McDaniel SH, Grus CL, Cubic BA, Hunter CL, Kearney LK, Schuman CC, ... & Miller BF (2014). Competencies for psychology practice in primary care. *American Psychologist*, 69, 409–429. [PubMed: 24820690]

- Morris SB, & DeShon RP (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105–125. [PubMed: 11928886]
- O'Donnell AN, Williams M, & Kilbourne AM (2013). Overcoming roadblocks: Current and emerging reimbursement strategies for integrated mental health services in primary care. *Journal of General Internal Medicine*, 28, 1667–1672. [PubMed: 23733375]
- Olfson M, Kroenke K, Wang S, & Blanco C (2014). Trends in office-based mental health care provided by psychiatrists and primary care physicians. *The Journal of Clinical Psychiatry*, 75, 247–253. [PubMed: 24717378]
- Pantziaras I, Fors U, & Ekblad S (2015). Training with virtual patients in transcultural psychiatry: Do the learners actually learn? *Journal of Medical Internet Research*, 17(2), e46. doi: 10.2196/jmir.3497. [PubMed: 25689716]
- Parsons TD (2015). Virtual standardized patients for assessing the competencies of psychologists. Retrieved from <http://www.igi-global.com/chapter/virtual-standardized-patients-for-assessing-the-competencies-of-psychologists/113106>.
- Parsons TD, Kenny P, Ntuen CA, Pataki CS, Pato MT, Rizzo AA, . . . Sugar J (2008). Objective structured clinical interview training using a virtual human patient. *Studies in Health Technology and Informatics*, 132, 357–362. [PubMed: 18391321]
- Patient Protection and Affordable Care Act, 42 U.S.C. § 18001 et seq (2010). Retrieved from: <http://housedocs.house.gov/energycommerce/ppacon.pdf>
- Petracchi H, & Collins K (2006). Utilizing actors to simulate clients in social work student role plays: Does this approach have a place in social work education? *Journal of Teaching in Social Work*, 26, 223–233.
- Pinquart M, Juang LP, & Silbereisen RK (2003). Self-efficacy and successful school-to-work transition: A longitudinal study. *Journal of Vocational Behavior*, 63, 329–346.
- Pollard RQ Jr, Betts WR, Carroll JK, Waxmonsky JA, Barnett S, deGruy III FV, ... & Kellar-Guenther Y (2014). Integrating primary care and behavioral health with four special populations: Children with special needs, people with serious mental illness, refugees, and deaf people. *American Psychologist*, 69, 377–387. [PubMed: 24820687]
- Posel N, Mcgee JB, & Fleischer DM (2014). Twelve tips to support the development of clinical reasoning skills using virtual patient cases. *Medical Teacher*, 37, 813–818. [PubMed: 25523009]
- Primm AB, Vasquez MJT, Mays RA, Sammons-Posey D, McKnight-Eily LR, Presley-Cantrell LR, . . . Perry GS (2010). The role of public health in addressing racial and ethnic disparities in mental health and mental illness. *Preventing Chronic Disease: Public Health Research, Practice and Policy*, 7(1), 1–7.
- Rawlings MA (2012). Assessing BSW student direct practice skill using standardized clients and self-efficacy theory. *Journal of Social Work Education*, 48, 553–576.
- Rawlings M, Bogo M, Katz E, and Johnson B (2012, 10). Designing objective structured clinical examinations (OSCE) to assess social work student competencies. Paper session presented at the Council on Social Work Education 58th Annual Program Meeting - Social Work: A Capital Venture, Washington, DC.
- Reeves WC, Strine TW, Pratt LA, Thompson W, Ahluwalia I, Dhingra SS, ... & Morrow B (2011). Mental illness surveillance among adults in the United States. *Mental and Mortality Weekly Report*, 60 (3), 1–29. Retrieved from <http://origin.glb.cdc.gov/mmWR/preview/mmwrhtml/su6003a1.htm>.
- Reinsmith-Jones K, Kibbe S, Crayton T, & Campbell E (2015). Use of *Second Life* in social work education: Virtual world experiences and their effect on students. *Journal of Social Work Education*, 51, 90–108.
- Sampson M, Parrish DE Parrish DE & Washburn M (2017). Assessing MSW students integrated behavioral health skills using an Objective Structured Clinical Examination. *Journal of Social Work Education*. doi:10.1080/10437797.2017.1299064
- Satter RM, Cohen T, Ortiz P, Kahol K, Mackenzie J, Olson C, . . . Patel VL (2012). Avatar-based simulation in the evaluation of diagnosis & management of mental health disorders in primary care. *Journal of Biomedical Informatics*, 45, 1137–1150. [PubMed: 22884757]

- Singh T, & Rajput M (2006). Misdiagnosis of bipolar disorder. *Psychiatry*, 3, 57–63.
- Skovholt TM, & Rønnestad MH (2003). Struggles of the novice counselor and therapist. *Journal of Career Development*, 30, 45–58.
- Talbot TB, Sagae K, John B, & Rizzo AA (2012). Sorting out the virtual patient: How to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. *International Journal of Gaming and Computer-Mediated Simulations*, 4(3), 1–19.
- Thériault A, Gazzola N, & Richardson B (2009). Feelings of incompetence in novice therapists: Consequences, coping, and correctives. *Canadian Journal of Counselling*, 43, 105–119.
- Triola M, Feldman H, Kalet A, Zabar S, Kachur E, Gillespie C, . . . Lipkin M (2006). A randomized trial of teaching clinical skills using virtual and live standardized patients. *Journal of General Internal Medicine*, 21, 424–429. [PubMed: 16704382]
- Turner HM, & Bernard RM (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Science and Disorders*, 33, 42–55.
- Washburn M, Bordnick P & Rizzo AS (2016). A pilot feasibility study of virtual patient simulation to enhance social work students brief mental health assessment skills. *Social Work in Health Care*, 55(9), 675–693. [PubMed: 27552646]
- Washburn M & Zhou S (in press). Technology enhanced clinical simulations: Tools for practicing clinical skills in online social work programs. *Journal of Social Work Education*.
- Wilkening GL, Gannon JM, Ross C, Brennan JL, Fabian TJ, Marcsisin MJ, & Benedict NJ (2016). Evaluation of branched-narrative virtual patients for interprofessional education of psychiatry residents. *Academic Psychiatry*, 1–5. [PubMed: 27921265]
- Wayne J, Bogo M, & Raskin M (2010). Field education as the signature pedagogy of social work education. *Journal of Social Work Education*, 46, 327–339.
- World Health Organization, World Organization of National Colleges, & Academic Associations of General Practitioners/Family Physicians. (2008). Integrating mental health into primary care: a global perspective. Retrieved from http://apps.who.int/iris/bitstream/10665/43935/1/9789241563680_eng.pdf

Figure 1.
Study Design

Figure 2.
Virtual Patient Armando Castilla

Table 1.

Participant Demographic Information by Condition

	Total	VPP	VPN	SAP
Characteristic	<i>M (SD)</i> 22	<i>M (SD)</i> 7	<i>M (SD)</i> 8	<i>M (SD)</i> 7
Age	28.95 (6.71)	26.25 (2.93)	31.13 (9.95)	29.14 (4.33)
Program Year	2.18 (0.91)	2.14 (1.35)	2.13 (0.83)	2.29 (0.49)
Characteristic	<i>N (%)</i>	<i>N (%)</i>	<i>N (%)</i>	<i>N (%)</i>
Gender				
Male	4 (18.2)	2 (28.6)	1 (12.5)	1 (14.3)
Female	18 (81.8)	5 (71.4)	7 (87.5)	6 (85.7)
Ethnicity				
African American	5 (22.7)	1 (14.3)	1 (12.5)	3 (42.9)
Caucasian Non-Hispanic	11 (50.0)	3 (42.9)	5 (62.5)	3 (42.9)
Asian/Pacific Islander	2 (9.1)	1 (14.3)	1 (12.5)	0 (0.00)
Hispanic Latino (Any)	2 (13.6)	1 (14.3)	1 (12.5)	1 (14.3)
Native American	1 (4.5)	1 (14.3)	0 (0.00)	0 (0.00)

Note. VPP – virtual patient with practice; VPN – virtual patient with no practice; SAP – standard actor with practice; Program year – current year of graduate training

Table 2.

DRF Scores by Group

Group N	Pre <i>X</i> (<i>SD</i>)	Post <i>X</i> (<i>SD</i>)	Change <i>X</i> (<i>SD</i>)	<i>t</i>	<i>p</i>	<i>g</i>
SAP(7)	23.57 (3.78)	25.71 (6.07)	2.14 (8.59)	-.66	.53	.21
VPN (8)	19.63 (4.47)	20.88 (3.68)	1.25 (5.87)	-.60	.57	.18
VPP (8)	16.56 (4.14)	24.00 (5.35)	7.14 (7.56)	-2.50	.05 [*]	.85

Note. SAP = Standard Actor Patient with Practice, VPN = Virtual Patient with No Practice, VPP = Virtual Patient with Practice.

^{*}
= $p < .05$

Table 3.

Percentage of Correct Diagnoses by Group

Group	Completely Correct	Partially Correct	Completely Correct
SAP – Pre	0.0%	71.4%	28.6%
SAP - Post	0.0%	42.9%	57.1%
VPN – Pre	25.0%	62.5%	12.5%
VPN – Post	25.0%	50.0%	25.0%
VPP - Pre	28.6%	57.1%	14.3%
VPP - Post	0.0%	71.4%	28.6%

Note. SAP = Standard Actor Patient with Practice, VPN = Virtual Patient with No Practice, VPP = Virtual Patient with Practice

Table 4.

Self-Efficacy Scores by Group

Group N	Pre <i>X</i> (<i>SD</i>)	Post <i>X</i> (<i>SD</i>)	Change <i>X</i> (<i>SD</i>)	<i>t</i>	<i>p</i>	<i>g</i>
SAP (7)	65.74 (16.17)	79.07 (7.69)	13.56 (11.52)	-3.07	.02 *	1.54
VPN (8)	68.63 (7.37)	73.88 (11.57)	5.25 (11.57)	-1.28	.24	0.30
VPP (8)	66.56 (6.44)	80.29 (5.74)	13.43 (6.05)	-5.87	.01 *	2.15

Note. SAP = Standard Actor Patient with Practice, VPN = Virtual Patient with No Practice, VPP = Virtual Patient with Practice.

*
= $p < .05$

Table 5.

Usability Scores (UFF)

Question	<i>M</i>	<i>SD</i>
Use this training method frequently	3.00	.76
Training method unnecessarily complex	2.77	.69
Easy to use	3.23	.81
Needed the support of a technical person	2.77	.75
Functions were well integrated	2.90	.53
Too much inconsistency	2.45	.80
Most people would learn to use this quickly	3.23	.69
Awkward to use	2.27	.94
Confident using this training method	2.55	.60
Needed to learn a lot of things before I could use	2.90	.61
<i>Overall usability (Sum of Scores $\times$ 2.5)</i>	<i>70.23</i>	<i>12.50</i>