



2017

Using Standardized Patients for Training and Evaluating Medical Trainees in Behavioral Health

Jeffrey D. Shahidullah

Rutgers University and Robert Wood Johnson Medical School, New Brunswick, NJ

Paul W. Kettlewell

Geisinger Health System, Danville, PA

Follow this and additional works at: <https://dc.etsu.edu/ijhse>



Part of the [Interprofessional Education Commons](#), [Mental and Social Health Commons](#), and the [Mental Disorders Commons](#)

Recommended Citation

Shahidullah, Jeffrey D. and Kettlewell, Paul W. (2017) "Using Standardized Patients for Training and Evaluating Medical Trainees in Behavioral Health," *International Journal of Health Sciences Education*, 4(2).

Available at: <https://dc.etsu.edu/ijhse/vol4/iss2/3>

This Article is brought to you for free and open access by the Journals at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in International Journal of Health Sciences Education by an authorized editor of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Using Standardized Patients for Training and Evaluating Medical Trainees in Behavioral Health

Introduction

Training delivered to medical students and residents in behavioral health (BH) is widely acknowledged as inadequate (McMillan, Land, & Leslie, 2017). While the use of standardized patients (SP) is common in medical training and education for physical health conditions via the adherence to clinical protocols for evaluation and treatment, this approach is infrequently used for BH conditions (i.e., psychiatric disorders, psychosocial/lifestyle factors). The few published studies using SP encounters for trainees on BH issues include the topics of substance abuse (MacLeod, Hungerford, Dunn, & Hartzler, 2008; Wamsley et al., 2013; Wilk & Jensen, 2002), lifestyle factors associated with chronic disease management (Cohen, Kitai, David, & Ziv, 2014), dementia and depression identification in geriatric populations (Westmoreland, Counsell, Tu, Wu, & Litzelman, 2010), suicide risk assessment (Fallucco, Hanson, & Glowinski, 2010; Fallucco, Conlon, Gale, Constantino, & Glowinski, 2012), communication skills (Douglas et al., 2016), and self-efficacy enhancement (Jerant et al., 2016). Participants in these studies include residents from internal medicine (Wamsley, et al., 2013), psychiatry (Arbuckle et al., 2013), internal medicine-pediatrics (Westmoreland et al., 2010), pediatrics (Fallucco et al., 2010; Lane, Ziv, & Boulet, 1999), and family medicine (Brown, Doonan, & Shellenberger, 2005), as well as medical students (Vest et al., 2016).

Used under specific circumstances, SP encounters have long been considered a reliable method of training and assessing trainee performance on addressing physical health conditions, and are even comparable to ratings of directly observed encounters with real patients (Wass, Jones, & Van der Vleuten, 2001). Therefore, it is worthwhile to consider whether this approach can feasibly be used in training and evaluation for BH skills. The broad purpose of this article is to review key features and evidence associated with using SPs in BH. Some of the key limitations in current BH training and assessment approaches are identified. Current evidence regarding the use of SPs as part of BH training and assessment with medical students and residents will be considered as well as some of the challenges. Suggestions for ways to address those challenges are presented. Our perspective is that the use of SPs in BH training holds promise and further implementation and research is needed.

How are Medical Trainees Currently Trained in Behavioral Health?

The primary forms of medical training are lecture- and textbook-based didactics and generic exposure to patients in clinic settings. Compared to didactics (e.g., classroom lectures/readings, case conferences, psychosocial rounds), clinical exposure to patients around BH topics represents a highly acceptable form of learning that trainees share is most beneficial to their development (Hampton, Richardson, Bostwick, Ward, & Green, 2015). While structured training such as direct observation of trainees in clinical tasks, with subsequent performance feedback using structured protocols in specified clinical areas are included as a Liaison Committee on Medical Education accreditation standard (2014), they are infrequently conducted (Kim et al., 2016). It is especially noteworthy that most commonly, feedback on clinical performance tends to occur in non-standardized ways as feedback provided to trainees may not be structured to address a range of clinical abilities from conducting a physical exam to addressing psychosocial issues.

Although BH problems occur frequently and, therefore, trainees receive some exposure to BH problems in clinical training settings, the BH problems are generally not the focus of training experiences provided by their physician supervisors and are either not addressed or are of secondary concern. Trainees receive substantially less training related to BH problems than physical health problems (Larson, Kettlewell, & Shahidullah, 2016). Additionally, when training in BH does occur in medical school this exposure is largely based on observations and shadowing, and thus, is less experiential than one would hope. Medical students complete clerkships or other rotations in mental health settings, although these rotations often focus on patients with serious mental illness, such as schizophrenia, rather than more commonly occurring BH issues (e.g., ADHD, depression, anxiety, suicidality, disruptive/externalizing behaviors) in which they more likely are expected to address in clinical care. Also, these rotations are typically adult-based, not pediatric-based (Larson et al., 2016).

During medical residency, most training in BH comes through one or two rotations, with the exception of psychiatry residents. For example, pediatrics residents receive the Accreditation Council for Graduate Medical Education (ACGME)-mandated 4-week developmental-behavioral pediatrics rotation as their primary exposure to BH. Outside of this rotation, BH training occurs via isolated instances of patient care in which the quality of learning is dependent on the level of guidance and supervision received from attendings and precepting faculty. Unfortunately, residents report feeling the supervision they receive is suboptimal as they do not feel their attendings and preceptors themselves feel confident in managing these types of concerns (Hampton et al., 2015). This often leads residents to report feeling that they under-recognize BH conditions.

Regarding treatment, residents also report that they feel they overly rely on biologically based interventions, such as psychotropic medications, rather than psychosocial interventions (Larson et al., 2016). Residents report they generally struggle with establishing rapport with patients and have difficulties in communication and relationship-building with patients and families with BH issues (Larson et al., 2016). An increasing amount of literature (Foy, 2010) has emphasized the importance of trainees developing clinical competence in “Common Factors” associated with care delivery for BH issues. Common factors are trans-diagnostic approaches valuing interpersonal communication skills, empathy/listening skills, using person-centered language, partnering with families, reinforcing strengths, and motivational interviewing. These are skills that have been shown necessary to improve the clinical care of all patients, regardless of diagnosis (Wissow et al., 2008).

When medical trainees receive lecture-based didactic exposure on BH topics, additional limitations exist. Fallucco et al. (2010) found poor retention of knowledge from didactic-based lectures on BH topics, such as suicidality, when compared to more active learning approaches such as clinical exposure. Larson et al. (2016) found that residents often report that the lectures they receive are too oriented towards their preparation for their board exams, and not as directly applicable to clinical practice.

Both the descriptions of the current approaches in BH training (with limited experiences in BH for trainees and minimal attention to BH concerns of patients in training settings where physical health concerns are emphasized) as well as the feedback from trainees regarding their views that

BH training is inadequate provide ample evidence that improvements in BH training for medical students and residents are needed.

How are Medical Trainees Currently Assessed in Behavioral Health?

Due to the difficulty in obtaining competency data in managing BH concerns, assessment of these competencies often stems from trainee self-reports pertaining to their attitudes (i.e., comfort/confidence in managing these conditions), knowledge, or skills. However, other more objective tools exist for assessing training competencies in BH. Due to the unpredictable nature in which BH referral concerns present and the difficulty in conducting live observation, trainees may also be assessed on their response to case vignettes in which a trainee responds to a pre-specified clinical presentation or referral concern and lists the steps in carrying out evidence-based evaluation and treatment or makes an accurate diagnosis (Steele, Lochrie, & Roberts, 2010). However, these vignettes are not able to adequately assess patient communication/interpersonal skills or common factors.

Chart review of clinical care of patients with BH concerns may be used to assess outcomes like use of screening tools, identification rates, and referrals made for BH concerns. However, these training tools cannot assess and provide feedback on patient-provider interaction skills and whether the identification of a BH concern was valid and if the referral was clinically indicated. For assessment of live interaction with real patients, a commonly used tool is the mini-CEX (i.e., mini-clinical evaluation exercise; Kim et al., 2016) in which a trainee is directly observed over a short duration (e.g., typically 10-20 minutes) to obtain a “snapshot” of the trainee’s interactions, and to then receive feedback on their competencies in history taking, physical examination and patient interaction skills, for example. This is difficult to do in real-time as each patient is context-specific and there is often difficulty in ascertaining the validity of the raters’ scores unless there are multiple observers present. For BH conditions specifically, it is often difficult to measure the subjective nature of patient interactions and/or clinical performance. Due to these content-specificity problems, the mini-CEX presents limitations.

In addition to the methods described above, (i.e., trainees giving self-reports of BH attitudes, knowledge and skills; their responses to case vignettes; use of chart reviews to assess compliance with BH guidelines; and use of live observation of clinical encounters with real patients as ways of assessing BH skills), SPs have also been used in some research and clinical settings as a BH assessment approach. A common assessment tool that uses SPs is the Objective Structured Clinical Examination (OSCE) (Harden, Stevenson, Downie, & Wilson, 1975). OSCEs are used to measure clinical tasks as well as communication and interpersonal skills. Other names for OSCE include CPX (Clinical Performance Exam), PBA (Performance-based Assessment), or CSA (Clinical Skills Assessment). OSCEs are typically scored using case-specific scoring sheets with a maximum achievable point allotment for each case. OSCEs can be used in the context of medical trainees addressing SP concerns. Details about use of SPs in BH training and assessment will be presented in the sections that follow.

It is clear that assessing medical students and residents’ BH skills presents challenges. There is no “gold standard” regarding how assessment needs to occur and most would agree that multiple methods of assessment are needed. Each of the approaches have inherent limitations. More

widespread use of SPs in assessing BH skills may broaden the current approaches and improve the adequacy of assessment of trainees' BH competencies.

Key Features of Using Standardized Patients in Behavioral Health Training and Assessment

SPs are often professional actors or students/graduates of theater programs that are recruited and taught by medical training faculty to portray a scenario involving a clinical complaint. Medical trainees are instructed to perform routine tasks related to interviewing and assessment and are often tasked with making a diagnosis and/or adhering to a clinical protocol. Although performed in a controlled setting, these encounters are intended to mimic real-world patient encounters. SPs are used to ensure consistent clinical scenarios with high degrees of reproducibility, which is facilitated, by using a standardized script. Medical training faculty, such as physicians or specialist physicians, develop the scenarios, often based on actual cases from clinic practices. SPs typically receive varying levels of training ranging from 4 hours (Fallucco et al., 2010) to 2-3 days (Cohen et al., 2014).

Many training programs will have a medical simulation or standardized patient center that specifically focuses on developing these types of opportunities within the site. Training faculty in conjunction with the hospital's standardized patient committee, for example, work collaboratively to create cases that are realistic and contain enough scripted details for the SP to present a problem/referral concern(s), medical and social history, and general social and emotional tone. The Association of Standardized Patient Educators (ASPE, 2009) has outlined generic suggestions for developing SP cases. The scenarios are often enacted in clinic exam rooms for setting authenticity. These encounters may occur in the context of *announced* (i.e., trainee made aware of upcoming SP encounter in advance) or *unannounced* (i.e., SP encounter occurs without trainee's foreknowledge) clinic visits. Two-way mirrors or video/audio equipment are typically used to allow unobtrusive observation. When SP encounters are videotaped, this allows for targeted performance feedback to be provided to the trainee as they can watch themselves in the encounter which may provide an enhanced learning experience.

Standardized Patients as a Training Tool

Vest et al. (2016) found the longitudinal use of a series of SP encounters to train family medicine clerkship students led to increased confidence in establishing relationships and in managing patients with chronic conditions, such as lifestyle counseling. Cohen et al. (2014) found that a combination of lectures with SP simulations was more effective than lectures only in training physicians to coach patients to improve their self-management of asthma using lifestyle and environmental changes and self-monitoring, and adherence to treatment regimens. In a study using SP encounters with performance feedback, Westmoreland et al. (2010) found that trainees being taught on SP simulations demonstrated small improvements in clinical testing protocols. However, not all studies demonstrate that use of SPs as a training tool leads to improved outcomes. For example, Herbstreit et al. (2017) found that medical students who received training in managing medical crises using SPs did not show increases in medical knowledge when compared to medical students who only received training using a traditional didactic teaching approach.

In order to make more definitive conclusions about the effectiveness of using SPs in medical education, methodological rigor in the research is needed. A noteworthy weakness in much of the medical trainee research using SPs is the dearth of studies employing a randomized controlled design in assessing training effectiveness. Also, most of the studies use small sample sizes resulting in limited statistical power to assess whether the intervention worked. It is unknown whether the relatively modest treatment effects in these studies would demonstrate more robust findings if a more rigorous experimental design were used. Another limitation regarding the effectiveness of SPs as a training tool pertains to how the results in the studies are typically evaluated. Often, outcome assessments include dependent variables such as mental health symptom ratings (e.g., Strengths and Difficulties Questionnaire) (Wissow et al., 2008) or trainee self-report on attitudes, knowledge, and practices (Fallucco et al., 2012). However, assessment tools relying on trainee self-report of their confidence and competence in evaluating and treating BH conditions are subject to the Dunning-Kruger Effect bias (Hodges, Regehr, & Martin, 2001). This bias is common in the medical education literature and occurs when trainee's confidence does not match their actual competence, and occurs more frequently on topics in which trainees receive less training and exposure, such as BH.

Acceptability of Training Approach

Studies using SPs have inquired about residents' satisfaction with SP training for BH topics. Wamsley et al., (2013) study focusing on substance abuse identification and treatment asked residents about the strengths/weaknesses of a SP learning tool. Residents had mixed feelings regarding the use of SP assessments to teach clinical skills and the overall value of the SP training experience and whether they would recommend the approach to other trainees. Open-ended responses also generated mixed responses regarding the strengths/weaknesses of SP training. For example, some residents reported they enjoyed the ability to see their skills improving longitudinally from the pre- to post-assessment. Residents also found a strength to be the ability to practice skills in a safe environment with realistic cases, which may be low incidence in regards to what they would typically treat in the real world. However, some residents in the study reported a weakness being that the SP felt artificial and contrived. Related to this, residents recommended that different SP cases be used for the pre- and post-assessments. Arbuckle et al. (2013) surveyed residents and faculty supervisors about their attitudes and perceived barriers toward the use of SPs as a training tool for improving clinical care for patients with BH diagnoses (e.g., depression). Raters generally agreed that SP training could help with symptom monitoring, diagnostic and treatment plan formulation, patient-centered communication skills, and engaging the patient in improved self-management of symptoms. The SP approach is also well-liked by medical students/clerkship students, particularly when SP encounters are used longitudinally whereby medical students are exposed to a SP over a series of encounters (Vest et al., 2016). This allows medical students to experience the process and benefits of longitudinal care in simulating the ability to following a patient over months and years that they would not normally be able to do in the given time for a clerkship.

Strengths/Limitations of Using Standardized Patients as a Training Tool

The use of SPs represents an effective form of training that is viewed as acceptable by medical trainees. The standardized and controlled approach allows for equal access for all trainees in a training program to various clinical presentations, which can be scheduled at the convenience of the trainee and used at any location. SPs facilitate structured observations and feedback to be provided by medical faculty, but also allow for the actors themselves to provide feedback on other qualitative aspects of the interaction that may be less apparent to observers, such as demonstrating warmth and empathy via the use of body language, tone, and eye contact. Also, SP simulations facilitate a systematic approach to curricular activities in ensuring ample exposure for trainees to develop competencies on clinical problems that they may not receive in high volume otherwise, such as mental health or other psychosocial issues (e.g., cross-cultural factors, adherence/compliance issues). This training approach allows “teachable moments” to be created, rather than waited for. In addition, it creates a venue for trainees to address clinical presentations without the risk of causing harm to real patients, particularly in the context of trainee’s novice skills. This aligns with the medical ethical imperative of nonmaleficence (i.e., “do no harm” to the patient).

Trainees typically respond to SP encounters just as they would if they were actual patients. There is however, the possibility for trainees to adjust their behavior when they know a patient is used solely for training purposes. For example, when SPs are pre-announced, it is possible that trainees can demonstrate the targeted interviewing behaviors under prompted circumstances in which they know they are being evaluated. In some cases, it then becomes necessary to use unannounced clinic visits. However, when these unannounced visits occur in the real clinic setting and are put on a trainee’s patient list for the day there are sometimes occurrences which may still reveal that the patient is not real, such as lack of prior medical history in the medical record. Therefore, they must pose as “new” patients in the medical records systems. Also, in real clinic practice other health care providers are often expected to see a patient before the physician and begin to document information in the medical chart. This may be disruptive to clinic flow to have to explain to nursing staff for example that a patient is being created for training purposes. In addition, Westmoreland et al. (2010) discuss an unintended barrier to the use of unannounced SP visits in a continuity clinic for medicine-pediatrics residents in the context of a training study to assess early detection of signs/symptoms of dementia and depression. As part of conducting the routine physical exam for the SP, the SP may have a true health problem, such as high blood pressure, that is identified in the course of the encounter and addressed by the trainee, thus essentially de-railing the original purpose of the training encounter. Additionally, the general concern about using “deception” in BH research (as would be the case if presenting a SP as if it were a real patient) has been generally criticized as unacceptable unless the benefits of the research outweigh the drawbacks of using deception.

Other difficulties exist in conducting SP training, particularly when training child-focused providers. Due to ethical (e.g., potential exposure to emotional distress) and practicality reasons (e.g., being a SP requires emotional maturity and at least average/above average cognitive abilities), there is a general difficulty in getting access to trained child actors, particularly for BH concerns (Tsai, 2004). While children are used for clinical encounters addressing physical concerns (Lane, Ziv, & Boulet, 1999), they are seldom used for BH concerns. One notable

exception was a pilot program which used child and adolescent actors (ages 9-19 years old) as SPs in training residents and medical students in communication skills addressing mental health issues including anorexia nervosa, depression, separation anxiety, and ADHD (Brown et al., 2005). More commonly, adult actors are used to portray children and adolescents. For example, Fallucco et al. (2010) aimed to teach pediatric residents to assess adolescent suicide risk with a standardized patient module, but rather than use actual adolescents in the training, used “adolescent-appearing, adults” which ranged in age from 21-26 years. This issue poses larger questions related to the applicability of using SPs in training pediatric providers in BH issues, particularly in the context of a Task Force Report from the American Academy of Pediatrics (2009) highlighting the need for improved BH competencies from pediatric providers in addressing common conditions such as ADHD, anxiety, depression, and suicidality. A final negative of training using SPs include that they can be costly for a training program when considering additional time and human resources pertaining to hiring and training human actors. For example, one internal medicine residency training program (Wamsley et al., 2013) estimated over 500 person-hours necessary for them to develop and implement the SP program, with a large proportion of these costs related to time allocated to developing and refining just three SP cases. In many instances, having more than three different cases will be ideal to use in training in order to create diverse representations of patients/referral concerns. For many programs, the training benefit may not outweigh these costs. It is difficult to determine whether any improvement gleaned from SPs are quantifiably meaningful as a justification for increased training costs.

Standardized Patients as an Assessment Tool

Considering the inherent limitations of current methods to medical trainee assessment in BH, SP simulations present promising solutions. Despite SPs having been used in the United States for medical education since the 1960s (Barrows & Abrahamson, 1964) many in the medical field believe they are underutilized in contemporary training curricula (Weaver & Erby, 2012). Traditionally used for physical health concerns (Buellens, Rethans, Goedhuys, & Buntinx, 1997), there have been comparatively limited studies using SPs in the context of BH assessment. SP simulation may be used as a tool to assess the effectiveness of another training tool (Jerant et al., 2009; Wissow et al., 2011). An example of this is a program using pre-/post-training SP encounters in which trainees are rated on their communication skills and patient-centered interactions with patients presenting with BH problems after exposure to a brief communication skills training (Wissow et al., 2011). Another example is examining mean change scores on performance with a SP encounter on identifying signs and symptoms of BH problems (e.g., dementia, depression) both before and after exposure to training (Westmoreland et al., 2010). Training faculty may develop an assessment checklist with items measuring a resident’s performance concerning pre-specified objectives/competencies. Checklist items can be dichotomous (e.g., “yes” or “no”) regarding whether a task/skill was completed, or use multi-point rating scales (e.g., “outstanding” to “unacceptable”; “accomplished”, “partially accomplished”, or “not accomplished”).

Use of SPs as part of assessing BH competencies require at least some minimal methodological rigor. If SPs are used as an assessment measure to ascertain the degree of training effectiveness for different training modalities or dosages for example, it is necessary for evaluators to be blind

or unaware of the trainees group allocation to a training group so that bias does not interfere. Also, if more than one evaluator is used to rate the performance, then a measurement of interrater reliability is necessary. When used for research, researchers may not only need to gather information on trainee's performance, but also on the accuracy of the SP in the case portrayal (i.e., presentation of essential clinical/interpersonal features presented correctly in each encounter). This is infrequently conducted in published SP training studies, although often training programs that wish to use SPs may have an accuracy requirement that SPs are required to meet before they are allowed to portray a case in a training encounter with an actual trainee. Assessments pertaining to reliability (i.e., degree of accurate measurement) should be obtained including intra-SP reliability (i.e., consistency of the SP presentation on different occasions) and inter-SP reliability (i.e., consistency across presentations from different SPs for the same case). Assessments of validity (i.e., degree to which a test measures what it intends to measure) should also be included.

When using SPs for research, there is typically a distinction between using *direct vs. indirect* methods for information gathering (Beullens et al., 1997). Direct methods include an observer, often a researcher, observing a trainee in delivering clinical care to a patient and using audio, video, or observation through a two-way mirror, or ratings directly from the SPs themselves. Indirect methods may be used when direct observation of a patient-provider interaction is not possible or when other outcome metrics are sought such as medical chart abstraction/audits looking at variables such diagnosis and referral rates. However, these indirect methods cannot assess outcomes such non-verbal communication and other common factors that are more suited to direct observation. It is noteworthy that when medical trainees know they are being observed in the context of a SP, they may act differently because of the *Hawthorne Effect*, i.e., the alteration of behavior by the subjects of a study due to their awareness of being observed (Paradis & Sutkin, 2017).

From a research perspective, trainees may typically complete a SP encounter at two or more time points and receive exposure to some type of training during the interim. However, inferences of causality may be limited in training studies in medical education that simply use a before/after design, as it may be difficult to demonstrate that all other aspects of the training curricula are controlled for (e.g., all trainees may not conduct the same rotations or in the same order). Also from a research perspective, the use of SPs may be used to document pre-vs post-performance ratings in response to some type of intervention received by residents to improve performance. Studies that use a control group in which to compare these innovations in training are not common due to the relatively small resident class sizes in which to allocate residents to different training groups. One way to increase the overall sample size in a study is to identify comparable training programs based on demographic and training/exposure variables as well as a baseline scores on an indicated measure.

However, when using this approach to research related to a training effectiveness study, it is difficult to use SPs, as there are often restrictions/barriers to their ability to travel to other training sites to perform SP encounters and certainly for residents at other training sites to travel to the location of the SP. Also, because of the difficulty in ensuring consistency in SP clinical scenarios across training sites, replication studies at different sites are rarely conducted. Without

studies being replicated by other programs, it is difficult for training curricula to demonstrate that its outcomes are generalizable.

Strengths/Limitations of Using Standardized Patients as an Assessment Tool

Studies using SP simulations as evaluation tools show a limited range of constructs in which they are suited to measure, often narrowing their focus to communication skills/common factors assessment. While some studies have used SP simulations to assess adherence to clinical evaluation and diagnostic protocols for suicide risk assessment or recognizing signs and symptoms of depression or dementia (Westmoreland et al., 2010), these evaluations are not generally viewed as a reliable tool for training to improve accuracy in diagnosing psychiatric disorders. This is likely due to the extensive training involved for the SP to properly grasp the subjective and contextual nature of the conditions as well as the social, developmental, behavioral and cultural factors, which affect the portrayal, compared to physical health conditions. Extensive selection and training of SPs is required to obtain sufficient reliability and validity of the SP clinical portrayals and consistency across trainees in which they interact. Trainee encounters with the SP is typically limited to one encounter, when in the real world, diagnostic and therapeutic interventions are often spread over the course of several encounters (Beullens et al. 1997). Because of the time-intensive nature of the SP training program and the fixed and limited number of trainees who can participate at each site, this can create difficulty in conducting validity and reliability analyses for checklist items. Also, for many training programs, video/audio recording or even use of two-way mirrors in SP encounters may be unfeasible due to resource limitations.

Ideally, conducting OSCEs with SPs should offer the same conditions to all trainees in the study to compare relativity in response to the same stimuli. Unfortunately, it is difficult to control for this with actors, as the nuances of each encounter (e.g., different words, conversation, tones, and emotional responses) will likely predispose varying degrees of deviation from the original role script. Therefore, the training that is involved in ensuring actors be as accurate and close to the standardized script as possible can be extensive. A line of research has evolved focusing specifically on the aspect of training SP actors including methods such as using video in SP training to improve portrayal accuracy (Schlegel, Bonvin, Rethans, & van der Vleuten, 2015). For programs eager to implement SP simulations, the resources necessary to allocate to ensuring high-quality evaluation may be prohibitive.

The difficulty in demonstrating consistency across encounters, providers, and SPs is a limitation posed to using SPs as an assessment tool. One way in which programs have sought to standardize the SP encounter is in using virtual patients (VPs) rather than live patients. A VP is an interactive computer-based simulation of a real-life clinic encounter. VPs present a standardized way of ensuring all trainees receive access to the same scenario and this also facilitates use of multi-site research or distance-learning environments as the VP can be used anywhere, thus remedying the problem of reduced trainee access to real patients. The trainee is expected to make judgments and clinical decisions based on the VPs presentation and response, the same way they would with a live SP (Ellaway, Poulton, Fors, McGee, & Albright, 2008). Problems with the use of VPs is that they can be complicated and costly to develop. Also, they are best suited to patient activities around solving a problem, such diagnosis/treatment, and

learning clinical process and thinking skills. However, these computer-based encounters may be less suited to developing patient-centeredness skills and use of common factors given the lack of human responsiveness to non-verbals, warmth, empathy, and understanding.

Summary: Potential Advantages and Disadvantages of Using Standardized Patients as Training and Assessment Tools in Behavioral Health

Advantages of using SPs in training include their availability, reliability, and controllability when properly trained, minimal risk to real patients, and the ability for trainees to receive structured performance feedback on infrequently occurring events. The process of using SPs in BH training is by definition “standardized” and thus more controlled. Trainee encounters with SPs can be scheduled, scripted, and repeated. This allows the “teachable moments” to be created rather than waiting for them to occur.

One could especially argue that use of SPs appears to be conceptually sound for training and assessing communication skills and common factors. This is based on the observation that other methods such as lectures, readings or didactics have not demonstrated to improve those skills. In addition to knowledge of evidence-based evaluation and treatment practice parameters, the essential competencies necessary for providing effective BH care reflect a set of skills that may be difficult to measure absent of observation of the patient-provider interaction context (e.g., communication skills, patient-centeredness, decision-making skills, self-efficacy and partnership enhancement, and cultural sensitivity). Therefore, SPs may offer a reliable way to teach and evaluate more comprehensive communication skills or common factors. Ideally, SPs would be used as a supplement (not in lieu of) to didactics and generic clinical exposure as this approach offers opportunities for trainees to practice skills in safe environments and particularly with low-incidence issues that they otherwise would receive little structured formal training and evaluation feedback.

Another advantage of using SPs as assessment tools for BH training include their ability for various training modalities to demonstrate their relative effectiveness on actual patient interactions compared to simply relying on trainee self-report of their attitudes and knowledge. This “show-how” vs “know-how” distinction is important as training programs are increasingly faced with finding ways to objectively measure competency in a number of areas defined by ACGME. In response to the need to assess clinical skills and performance, the use of SPs present a promising method of doing this.

Assessment of BH skills clearly is an area of study without a clear “gold standard”. There is no single or definitive way to measure BH skills of trainees that is widely accepted. This is true for many areas of research that require behavioral or psychological assessment. The primary response from researchers to this challenge has been to find multiple measurement approaches and conduct research about reliability and validity for each of those measurement methods as well as to conduct correlational studies among measures. Adding SPs as an additional tool represents a reasonable way to expand our ability to assess BH competencies in medical students and residents.

There are potential disadvantages and challenges when using SPs in BH training and assessment. First, trainees may view the use of SPs as somewhat artificial and contrived and the skills they might demonstrate in training situations with SPs may not convert to actual changes in skills when working with real patients. There is limited evidence that the BH skills demonstrated using SPs for BH training are generalizable outside of a particular training program and more research is needed. A second disadvantage of using SPs for training or assessment purposes includes the financial expense and the administrative time in recruiting, training, and developing clinical scenarios that are valid and reliable. Many training programs may not have the capacity in these areas that is necessary to initiate a SP training program.

Approaches to Addressing Challenges in Using Standardized Patients

How could we best respond to the awareness that using SPs may be perceived as artificial or contrived? First, we should simply acknowledge that possibility and then work to make the experience with SPs most closely resemble that of the real clinic setting. Scripts can be developed that best mirror how patients actually behave and the problems they present. Skilled actors can be used. The experience with SPs can be held in real clinic settings. Trainees can provide feedback about how adequately their experience represent real patient contact and improvements in the process can occur based on that iterative feedback from trainees. Using SPs for BH training in a thorough and detailed manner may lead to greater generalizability of skills to real world settings.

What about the amount of effort, resources and costs associated with developing the scripts and the preparation and details needed to adequately use SPs in BH training and assessment? The entire process of using SPs well cannot likely be done cheaply or easily and that needs to be acknowledged. A discussion about how we make decisions about using health care resources for patient care and training of physicians is well beyond the scope of this article. However, priorities for training of health care providers, including BH training, deserve discussion among national professional groups such as the American Council of Graduate Medical Education. Also, these debates need to be held locally in academic medical centers and integrated health care organizations to address how limited resources should be spent. Advocates for BH training and research are needed.

Conclusion

Primarily because medical students and residents need more extensive BH training and because current BH training models and BH assessment approaches are inadequate, we believe that innovative models of BH training and assessment are needed. Use of SPs holds promise as one component of expanded and higher quality BH training and improved assessment of BH competencies.

References

- American Academy of Pediatrics. (2009). Policy statement — The future of pediatrics: Mental health competencies for the care of children and adolescents in primary care settings. *Pediatrics*, *124*, 410–421.
- Arbuckle, M. R., Weinberg, M., Harding, K. J., Isaacs, A. J., Covell, N. H., Cabaniss, D. L., ... Sederer, L. I. (2013). The feasibility of standardized patient assessments as a best practice in an academic training program. *Psychiatric Services*, *64*, 209-211.
- Association of Standardized Patient Educators. (2009). Recommended standardized patient case outline. Retrieved February 23, 2017 from [http://aspeducators.org/img/mededportal%submit_criteria\(1\).pdf](http://aspeducators.org/img/mededportal%submit_criteria(1).pdf)
- Barrows, H. S., & Abrahamson, S. (1964). The programmed patient: A technique for appraising student performance in clinical neurology. *Journal of Medical Education*, *39*, 802-805.
- Brown, R., Doonan, S., & Shellenberger, S. (2005). Using children as simulated patients in communication training for residents and medical students: A pilot program. *Academic Medicine*, *80*, 1114-1120.
- Buellens, J., Rethans, J. J., Goedhuys, J., & Buntinx, F. (1997). The use of standardized patients in the research in general practice. *Family Practice*, *14*, 58-62.
- Cohen, A. G., Kitai, E., David, S. B., & Ziv, A. (2014). Standardized patient-based simulation training as a tool to improve the management of chronic disease. *Simulation in Healthcare*, *9*, 40-47.
- Douglas, S. R., de Andade, V., Boyd, S., Leslie, M., Webb, L., Davis, L., ..., Bickman, L. (2016). Communication training improves patient-centered provider behavior and screening for soldiers' mental health concerns. *Patient Counseling and Education*, *99*, 1203-1212.
- Ellaway, R., Poulton, T., Fors U., McGee, J. B., & Albright, S. (2008). Building a virtual patient commons. *Medical Teacher*, *30*, 170-174.
- Fallucco, E. M., Hanson, M. D., Glowinski, A. L., (2010). Teaching pediatric residents to assess adolescent suicide risk with a standardized patient module. *Pediatrics*, *125*, 953-959.
- Fallucco, E. M., Conlon, M. K., Gale, G, Constantino, J. N., & Glowinski, A. L. (2012). Use of a standardized patient paradigm to enhance proficiency in risk assessment for adolescent depression and suicide. *Journal of Adolescent Health*, *51*, 66-72.
- Foy, J. M. (2010). Enhancing pediatric mental health care: Algorithms for primary care. *Pediatrics*, *125*, S109-S125.

- Hampton, E., Richardson, J. E., Bostwick, S., Ward, M. J., & Green, C. (2015). The current and ideal state of mental health training: Pediatric resident perspectives. *Teaching and Learning in Medicine, 27*, 147-154.
- Harden, R., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using objective structured clinical examination. *British Medical Journal, 1*, 447-451.
- Herbstreit, F., Merse, S., Schnell, R., Noack, M., Dirkman, D., Besuch, A., & Peters, J. (2017). Impact of standardized patients on the training of medical students to manage emergencies. *Medicine, 96*, 1-5.
- Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: Novice physicians who are unskilled and unaware of it. *Academic Medicine, 76*, 87-89.
- Jerant, A. F., Kravitz, R. L., Tancredi, D., Paterniti, D. A., White, L., Baker-Nauman, L., ... Franks, P. (2016). Training primary care physicians to employ self-efficacy-enhancing interviewing techniques: Randomized controlled trial of a standardized patient intervention. *Journal of General Internal Medicine, 31*, 716-722.
- Jerant, A. F., Kravitz, R. L., Azari, R., White, L., Garcia, J. A., Virata, M. C., & Franks, P. (2009). Training residents to employ self-efficacy enhancing interviewing techniques: Randomized controlled trial of a standardized patient intervention. *Journal of General Internal Medicine, 24*, 606-613.
- Kim, S., Willett, L. R., Noveck, H., Patel, M. S., Walker, J. A., & Terregino, C. A. (2016). Implementation of a mini-CEX requirement across all third-year clerkships. *Teaching and Learning in Medicine, 28*, 424-431.
- Lane, J. L., Ziv, A., & Boulet, J. R. (1999). A pediatric clinical skills assessment using children as standardized patients. *Archives of Pediatric Adolescent Medicine, 153*, 637-644.
- Larson, S. L., Kettlewell, P. W., & Shahidullah, J. D. (2016). *As a pediatrician, I don't know the second, third, or fourth thing to do: Behavioral health and residency training*. Poster presented at the International Congress of Behavioral Medicine, Melbourne, Australia.
- Liaison Committee on Medical Education. (2014). *Functions and structure of a medical school: Standards for accreditation of medical education programs leading to the MD degree*. Retrieved February 10, 2017 from <http://www.lcme.org/publications.htm>.
- MacLeod, J. B., Hungerford, D. W., Dunn, C., & Hartzler, B. (2008). Evaluation of training of surgery interns to perform brief alcohol interventions for trauma patients. *Journal of American College of Surgeons, 207*, 639-645.

- McMillan, J. A., Land, M., & Leslie, L. K. (2017). Pediatric residency education and the behavioral and mental health crisis: A call to action. *Pediatrics*, *139*, 21-41.
- Paradis, E., & Sutkin, G. (2017). Beyond a good story: From Hawthorne Effect to reactivity in health professions education research. *Medical Education*, *51*, 31-39.
- Schlegel, C., Bonvin, R., Rethans, J. J., & van der Vleuten, C. (2015). The use of video in standardized patient training to improve portrayal accuracy: A randomized post-test control group study. *Medical Teacher*, *37*, 730-737.
- Steele, M. M., Lochrie, A. S., & Roberts, M. C. (2010). Physician identification and management of psychosocial problems in primary care. *Journal of Clinical Psychology in Medical Settings*, *17*, 103-115.
- Tsai, T. C. (2004). Using children as standardized patients for assessing clinical competence in paediatrics. *Archives of the Disabled Child*, *89*, 1117-1120.
- Vest, B. M., Lynch, A., McGuigan, D., Servoss, T., Zinnerstrom, K., & Symons, A. B. (2016). Using standardized patient encounters to teach longitudinal continuity of care in a family medicine clerkship. *BMC Medical Education*, *16*, 208.
- Wamsley, M. A., Julian, K. A., O'Sullivan, P., Satterfield, J. M., Satre, D. D., McCance-Katz, E., & Batki, S. L. (2013). Deigning standardized patient assessments to measure SBIRT skills for residents. A literature review and case study. *Journal of Alcohol and Drug Addition*, *57*, 46-65.
- Wass, V., Jones, R., & Van der Vleuten, C. (2001). Standardized or real patients to test clinical competence? The long case revisited. *Medical Education*, *35*, 21-25.
- Weaver, M., & Erby, L. (2012). Standardized patients: A promising tool for health education and health promotion. *Health Promotion Practices*, *13*, 169-174.
- Westmoreland, G. R., Counsell, S. R., Tu, W., Wu, J., & Litzelman, D. K. (2010). Web-based training in geriatrics for medical residents: A randomized controlled trial using standardized patients to assess outcomes. *Journal of the American Geriatrics Society*, *58*, 1163-1169.
- Wilk, A.I., & Jensen, N.M. (2002). Investigation of a brief teaching encounter using standardized patients. *Journal of General Internal Medicine*, *17*, 356-360.
- Wissow, L., Anthony, B., Brown, J., DosReis, S., Gadowski, A., Ginsburg, G., & Riddle, M. (2008). A common factors approach to improving the mental health capacity of pediatric primary care. *Administrative Policy in Mental Health*, *35*, 305-318.

Wissow, L., Gadowski, A., Roter, D., Larson, S., Lewis, B., & Brown, J. (2011). Aspects of mental health communication skills training that predict parent and child outcomes in pediatric primary care. *Patient Education and Counseling*, 82, 226-232.