Objective Structured Clinical Examination (OSCE) in Psychiatry Education: A Review of Its Role in Competency-Based Assessment

Christos Plakiotis

Abstract Over the last two decades, Objective Structured Clinical Examination (OSCE) has become an increasingly important part of psychiatry education and assessment in the Australian context. A reappraisal of the evidence base regarding the use of OSCE in psychiatry is therefore timely. This paper reviews the literature regarding the use of OSCE as an assessment tool in both undergraduate and postgraduate psychiatry training settings. Suitable articles were identified using the search terms 'psychiatry AND OSCE' in the ERIC (educational) and PubMed (healthcare) databases and grouped according to their predominant focus: (1) the validity of OSCEs in psychiatry; (2) candidate preparation and other factors impacting on performance; and (3) special topics. The literature suggests that the OSCE has been widely adopted in psychiatry education, as a valid and reliable method of assessing psychiatric competencies that is acceptable to both learners and teachers alike. The limited evidence base regarding its validity for postgraduate psychiatry examinations suggests that more research is needed in this domain. Despite any shortcomings, OSCEs are currently ubiquitous in all areas of undergraduate and postgraduate medicine and proposing a better alternative for competency-based assessment is difficult. A critical question is whether OSCE is sufficient on its own to assess high-level consultancy skills, and aspects of professionalism and ethical practice, that are essential for effective specialist practice, or whether it needs to be supplemented by additional testing modalities.

Keywords Objective structured clinical examination • OSCE • Psychiatry education • Medical education • Competency-based assessment

C. Plakiotis (✉)
Monash Ageing Research Centre (MONARC), Monash University, Melbourne, VIC, Australia

Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Melbourne, VIC, Australia

Aged Persons Mental Health Service, Monash Health, Melbourne, VIC, Australia

Aged Psychiatry Academic Unit (Monash University), c/o Medical Administration, Kingston Centre, Warrigal Road, Cheltenham, VIC, 3192, Australia
e-mail: Chris.Plakiotis@monash.edu

© Springer International Publishing AG 2017

P. Vlamos (ed.), *GeNeDis 2016*, Advances in Experimental Medicine and Biology 988, DOI 10.1007/978-3-319-56246-9_13

1 Background

Clinical competence, and its reliable and valid assessment, has increasingly become a focus for medical educators and a range of stakeholders with legitimate interests in the clinical competence of medical students and postgraduate medical trainees. In this context, it is essential that methods for assessing clinical competence effectively discriminate between adequate and inadequate performers [1].

The field of competency assessment in medical education was revolutionised in the 1970s with the introduction of the first objective structured clinical examination (OSCE), which evaluated the performance of professional behaviours through actors and choreographed scenarios [2, 3]. Candidates involved in OSCE rotate sequentially through several 'stations' at which specific tasks are undertaken in relation to structured cases. Tasks are usually clinically-focussed, e.g. taking a history or performing a physical examination or procedural skill. A specified time is allocated for each station and marking schemes are structured and pre-determined. OSCEs may vary with regard to the time allocated to individual stations, the use of mannequins or standardised patients, the choice of examiners (e.g. clinicians or standardised patients), and the marking method used (e.g. checklist or rating scale). However, the notion that all candidates must undertake the same task in the same amount of time and be marked using identical criteria is essential to all OSCEs [1].

In a seminal paper in the field of psychiatry education, Hodges [4] provided a comprehensive guide for developing, monitoring and improving the quality of psychiatry OSCEs that involved the following key elements:

1. Planning (budgeting, funding, team set-up, sourcing standardised patients)
2. Creation (blueprinting, station development, measurement instruments)
3. Preparation (recruiting and training standardised patients and examiners, venue location)
4. Conduct (preparation, useful documents, reporting results)
5. Quality control (monitoring and improvement, data analysis, standard setting)
6. Research.

In a commentary on the OSCE guide developed by Hodges [4], Davis [5] argued that medical education had stalled educationally and depended on 'attestation' as the means for assuring competence and quality for far too long. The focus on core competencies in medical education, and the use of OSCE to assess them, was identified as a potential driver of educational innovation. The scope of OSCEs to both educate and test was emphasised. A strategic selection of cases across the spectrum of competence was recommended to ensure the validity and reliability of the process. OSCEs' ability to evaluate performance in acute care situations (that would otherwise be practically and ethically difficult) and low-frequency but high-stakes situations (such as suicide risk assessment and management) was emphasised. However, a need to demonstrate a relationship between OSCE performance and clinical outcome was identified.

Over the last two decades, the OSCE has become an increasingly important part of psychiatry education and assessment in the Australian context. At an undergraduate (medical student) level, psychiatry stations frequently feature in OSCEs conducted jointly with other medical disciplines. The OSCE has also been a key component of the examination for Fellowship of the Royal Australian and New Zealand College of Psychiatrists. A review of the evidence base regarding the use of OSCE in psychiatry is therefore timely to evaluate the positive contributions it has made to psychiatry education and training as well as to consider its potential shortcomings.

2 Objectives

This paper reviews the literature regarding the use of OSCE as an assessment tool in both undergraduate and postgraduate psychiatry training settings. Suitable articles from English-speaking countries primarily were identified using the search terms 'psychiatry AND OSCE' in the ERIC (educational) and PubMed (healthcare) databases. Papers were grouped according to their predominant focus: (1) the validity of OSCEs in psychiatry; (2) candidate preparation and other factors impacting on performance; and (3) special topics. Key findings from the literature are subsequently discussed in light of the author's own experience as an undergraduate and postgraduate OSCE candidate and examiner. Given the increasing emphasis on interprofessional education and assessment for postgraduate health practitioners, the use of OSCE in the related fields of clinical psychology and neuropsychology is also briefly considered.

3 The Validity of OSCEs in Psychiatry

Papers focused predominantly on the validity of OSCEs in psychiatry are reviewed below and summarised in Table 1.

Loschen [6] reported on the use of OSCEs to assess the clinical skills of second- and fourth-year psychiatry residents at the Southern Illinois University School of Medicine since 1985. A separate examination was developed for each trainee group. Stations generally incorporated a written component. Each OSCE was comprised of six 40-min stations, allowing eight residents to complete it over 6 h. The OSCE format fared favourably as a method of formative assessment. However, marked differences between global OSCE ratings and clinical ratings of residents' performance by supervisors and unit heads prompted the authors to avoid its use for summative assessment until more acceptable performance standards could be established. A tendency to base OSCE scenarios on unusual or rare cases was discouraged as a marked departure from the goal of testing clinically relevant skills was felt to compromise the method's validity.

Table 1 Papers regarding the validity of OSCEs in psychiatry

Author year	Participants and OSCE characteristics	Key findings
Loschen [6]	• Second-year (18) and fourth-year (25) psychiatry residents • Six 40-min stations	• OSCE fared favourably as a formative assessment method • Differences between global OSCE ratings and clinical ratings of residents' performance by supervisors and unit heads prompted avoidance of its use for summative assessment, pending better performance standards • Scenarios based on unusual or rare cases are best avoided
Hodges et al. [7]	• Medical students (192) • Ten 12-min stations	• OSCE viably and reliably assessed high-level psychiatric skills in a format appreciated by students and teachers
Hodges and Lofchy [8]	• Medical students (42) • 'Mini-OSCE' (four 15-min stations)	• 'Mini-OSCE' demonstrated good interstation reliability and was well-received by students and academics • Benefits conferred without expense of longer OSCEs
Hodges et al. [9]	• Medical students (33) and psychiatry residents (17) • Eight 12-min stations	• Rankings on global scores, but not checklist scores, were accurately predicted by communication instructors • Performance was accurately predicted by supervisors (as measured by checklist scores but not global ratings) • Residents obtained significantly higher mean OSCE scores than medical students on global ratings only • Scenarios rated by residents as highly realistic
Loschen [10]	• Three stations for medical students (20–30-min simulated patient interview, 15–20-min testing) • Five to six stations for psychiatry residents	• OSCE was a cost-effective means of performance-based clinical skills assessment • Practical application of knowledge was evaluated (unlike traditional pen-and-paper knowledge-based tests)

Park et al. [12]	• Third-year medical students (286) • Nine-station psychiatry OSCE (including five 15-min standardised patient interviews)	• History taking, interpersonal skills and physical examination scores in a Clinical Skills Examination reflected scores for same skill set in an OSCE • Acceptable construct validity for both checklist and global process scores, supporting their continued use
Walters et al. [13]	• Fourth-year medical students (128) • Fifteen 'active' 6-min stations and two to three 'rest' stations	• Perceived face and content validity were high and overall reliability was moderate to good • OSCE was practically viable and well-received by students, simulated patients and examiners • Array of modalities enabled testing of many students and broad topic inclusion without cost of simulated patients
Marwaha [14]	• Debate paper	• OSCEs might be 'gold standard' in medical student clinical assessment, but their validity, authenticity and educational impact are open to challenge in postgraduate psychiatry • An inability to test complex real life scenarios and high-level psychiatric reasoning might compromise the standing of the profession and future patient care

Hodges et al. [7] assessed the feasibility, reliability, and validity of an OSCE for University of Toronto medical students undertaking their psychiatry rotation. A ten-station OSCE was developed in two parallel forms and administered three times. Performance was assessed via a checklist (content) and a global-rating scale (process) for each station. Additionally, each student's performance was graded as 'pass,' 'borderline' or 'fail' (by the examiner) and comments were recorded (by the examiner and standardised patient). Ninety-four students were examined with the first form (mean score 70.47%) and 98 with the second (mean score 67.66%). The findings suggested the viability and reliability of OSCE for assessing high-level psychiatric skills in a format appreciated by students and teachers.

In an effort to encourage psychiatry educators to adopt the OSCE in place of oral exams with inferior psychometric properties, Hodges and Lofchy [8] developed a four-station 'mini-OSCE' and trialled its use among 42 medical students undertaking their psychiatry clinical placement at the University of Toronto. Marks were assigned for content and process and the two domains equally weighted to derive a final score. Feedback was provided based on a global judgment of performance (as 'pass,' 'borderline,' or 'fail'). Exam scores ranged from 56 to 86%, with a mean of 74%. The mini-OSCE demonstrated good interstation reliability and was well-accepted by students and academic staff, benefits conferred without the expense associated with longer OSCE formats.

Hodges et al. [9] examined the validity of a psychiatry OSCE comprised of eight 12-min stations conducted among 33 medical students on their psychiatry rotation and 17 psychiatry residents at the University of Toronto. Medical students' rankings on global scores, but not checklist scores, were accurately predicted by communication instructors. Prediction of medical students' OSCE performance by faculty supervisors, as measured by checklist scores but not global ratings, was reasonably accurate. Residents obtained significantly higher mean OSCE scores than medical students on global ratings only. Scenarios were rated by residents as highly realistic. Based on these findings, the authors suggested the OSCE to be a valid tool for assessing medical students' clinical competence.

Loschen [10] described the role of OSCE in both resident evaluation and medical student teaching at the Southern Illinois University School of Medicine. This examination was composed of 5–6 stations and evolved over the years to include more comprehensive tasks than the circumscribed ones reported in an earlier paper [6]. A 3-station OSCE exam—after a format developed by Barrows et al. [11]—was used to test senior medical students' clinical skills competency at the conclusion of their psychiatry rotation. Each station involved a 20–30-min outpatient psychiatric evaluation of a standardised patient followed by 15–20-min of computerised short-answer and multiple choice questions. The author concluded that OSCE represented a cost-effective means of performance-based clinical skills assessment that allowed evaluation of the practical application of knowledge in a manner not possible with traditional pen-and-paper knowledge-based tests.

Park et al. [12] analysed aggregated archival data from two classes of third-year medical students (286 in total) at the St Louis University School of Medicine to assess the construct validity of the checklist and global process scores for an

OSCE in psychiatry. Binary checklists were used to evaluate clinical skills and numeric rating scales to assess standardised patients' perception of interpersonal skills. History taking, interpersonal skills and physical examination scores in a separate Clinical Skills Examination reflected scores for the same skill set in the psychiatry OSCE. The pattern of relationships that emerged confirmed acceptable construct validity for both the checklist score and global process score especially, supporting their continued use in psychiatry OSCEs.

Walters et al. [13] evaluated the validity and reliability of a multimodal OSCE developed on a limited budget for undergraduate psychiatry teaching at the Royal Free and University College Medical School in London. Four OSCEs were evaluated, comprised of 15–18 stations each and administered to 128 fourth-year medical students. Station types included: (1) history taking 'interactive' stations; (2) communication skills 'interactive' stations; (3) telephone communication with colleague; (4) video mental state examination; (5) written case vignette; and (6) written problem orientated vignette with visual prompt. Perceived face and content validity were high and overall reliability was moderate to good. The OSCE was practically viable and well-received by students, simulated patients and examiners. Integrating an array of modalities enabled the testing of many students and broad topic inclusion while avoiding the cost of simulated patient payments.

The Clinical Assessment of Skills and Competencies (CASC) is an OSCE used as a clinical examination gateway, granting access to becoming a senior psychiatrist in the United Kingdom (UK). In a debate paper, Marwaha [14] examined the utility of the CASC from the viewpoint of a senior psychiatrist. The author argued that while OSCEs might be the 'gold standard' in medical student clinical assessment, their validity, authenticity and educational impact are open to challenge in postgraduate professional psychiatry examinations. It was suggested that an inability to test complex real life scenarios and high level psychiatric reasoning might compromise the standing of the profession and future patient care.

4 Candidate Preparation and Other Factors Impacting on Performance

Papers regarding candidate preparation and other factors impacting on psychiatry OSCE performance are reviewed in this section and summarised in Table 2.

Blaskiewicz et al. [15] analysed aggregated archival OSCE performance data (in the form of a 25-item binary content checklist) from a class of 141 third-year medical students at Saint Louis University School of Medicine to investigate the impact of testing context and rotation order on student performance on a station common to both the obstetrics and gynaecology (O&G) and psychiatry OSCEs. Regardless of rotation order, students were less likely to address O&G issues in the psychiatry OSCE and psychiatric issues in the O&G OSCE. The authors concluded that testing context may bias students' collection and interpretation of patient

Table 2 Papers regarding candidate preparation and other factors impacting on performance

Author year	Participants and OSCE characteristics	Key findings
Blaskiewicz et al. [15]	• Third-year medical students (141) • OSCE station common to both O&G (1 of 10) and psychiatry (1 of 6) OSCEs	• Testing context may bias students' collection and interpretation of patient information • OSCEs may better simulate clinical reality by incorporating scenarios with diagnostic options not restricted to most recently completed rotation
Park et al. [16]	• Third-year medical students (869) • Psychiatry OSCE included five 15-min standardised patient interviews	• No association between performance and rotation timing • No association between performance trends and potential performance moderators (student preference for rotation order and specialty choice)
Robinson et al. [17]	• Psychiatry residents (26) • Monthly 2-h OSCE training session • Department-wide level OSCE (number of stations equal to year of training)	• Residents participating in OSCE training sessions (in which they authored and examined new stations) passed the psychiatry specialty examination of the Royal College of Physicians and Surgeons of Canada on their first attempt
Goisman et al. [18]	• Fourth-year medical students (136) • Nine stations (one 15-min psychiatry station)	• Psychiatry clerkship completion was associated with significantly better OSCE performance on phenomenology, mental state examination and differential diagnosis • Performance on history taking, communication skills and treatment planning was not affected
Morreale et al. [19]	• Medical students (active learning for 102 and standard lectures for 108) • Institutional-based OSCE (not graded)	• Active learning gave rise to superior student satisfaction and sense of preparedness for OSCE than standard lectures • Active learning included videotaped patient interviews, small group formulation of mental state examination and differential diagnosis, and discussion of these components and treatment planning with the clerkship director
Griswold et al. [20]	• Third-year medical students (integrated curriculum for 59 out of 922) • Comprehensive fourth-year OSCE (one psychiatry station)	• OSCE scores demonstrated effectiveness of an integrated curriculum incorporating didactic and clinical teaching, mentoring, psychotherapy experience, pharmacology instruction and immersive acute psychiatry exposure • Increased uptake of psychiatry as a career choice resulted

information and that OSCEs may better simulate clinical reality by incorporating scenarios with diagnostic options not restricted to the most recently completed rotation.

Park et al. [16] analysed archival data from six classes of third-year medical students (869 in total) at the St Louis University School of Medicine to see whether completion of the psychiatry rotation early in the clerkship cycle adversely affected student performance. No association was found between performance and rotation timing or between performance trends and potential performance moderators (student preference for rotation order and specialty choice).

Robinson et al. [17] described a training program to prepare psychiatry residents at the University of Western Ontario for the OSCE component of the Royal College of Physicians and Surgeons of Canada's (RCPSC) psychiatry specialty examination. Each month, residents attended a 2-h OSCE training session in which they acted as examiners for new stations they had authored themselves. A moderator facilitated the conduct of these training stations in a group setting. Once refined through a group feedback and discussion process, stations were incorporated in a bank of training stations. Residents also participated in a formal departmental OSCE twice a year. The OSCE training program was a key drawcard of the University's residency program and all 26 residents who completed it in the 5 years since its inception passed the RCPSC examination on their first attempt.

A psychiatry OSCE station completed by 136 fourth-year medical students at Harvard Medical School as part of a nine-station OSCE was used to compare the performance of students who had (46%) or had not (51%) completed a core psychiatry clerkship and determine which areas distinguished the two groups [18]. Psychiatry clerkship completion was associated with significantly better OSCE performance on phenomenology and mental state examination (six items) and differential diagnosis (three items). History taking, communication skills and treatment planning did not differ between groups.

Morreale et al. [19] demonstrated that, relative to conventional lectures, an active learning curriculum gave rise to superior student satisfaction and sense of preparedness for the OSCE at Wayne State University School of Medicine. Active learning seminars involved viewing video vignettes of patient interviews, formulating the mental state examination and differential diagnosis in small groups, and discussion of these components and treatment planning with the clerkship director.

A longitudinal, integrated, third-year psychiatry curriculum implemented within the Harvard Medical School-Cambridge Integrated Clerkship was demonstrated over 8 years to provide effective learning through OSCE scores, National Board of Medical Examiners shelf-exam scores, written work, and observed clinical work [20]. In addition to didactic and clinical teaching, the program incorporated longitudinal mentoring, psychotherapy experience, psychopharmacology instruction and immersive acute psychiatry exposure. An increased uptake of psychiatry as a career choice was a further benefit.

5 Special Topics

In this section, papers focussed on a range of special topics related to psychiatry OSCEs are reviewed. To further orient the reader, these papers are summarised in Table 3.

5.1 Attitudes of Learners and Teachers Towards Psychiatry OSCEs

Sauer et al. [21] evaluated the experience and views of psychiatry senior house officers in the UK on the inherent ability of OSCE to assess their clinical skills, using a modified version of the questionnaire developed by Hodges et al. [22] to assess Canadian psychiatry residents' attitudes to the OSCE. The authors conducted a mock OSCE based on Royal College of Psychiatrists' (RCPsych) OSCE format (twelve 7-min stations), but each station had three junior trainees (one candidate and two observers, 36 in total) and each candidate was examined at four stations instead of twelve. Trainees evaluated the OSCE positively, with 86% considering it fair and 89% appropriate in assessing clinical ability, and found it preferable to the individual patient assessment which it replaced in the RCPsych membership examination.

In a survey of 111 final-year medical students at University College Dublin, administered after the final psychiatry examination but prior to results being released, OSCE was the most highly-rated teaching and examination method [23]. The utility of an OSCE-based methodology in both teaching clinical skills and subsequently assessing the effectiveness of learning may have underpinned students' preference for this technique.

5.2 Standardised Patients in Psychiatry OSCEs

Sadeghi et al. [24] examined the views of 21 final-year psychiatry residents and 24 board-certified psychiatrist examiners regarding standardised patients' competence in simulating psychiatric disorders, and the accuracy of their performances, in an OSCE setting in Iran. The OSCE was comprised of eight 12-min stations and the standardised patients were university students or Tehran Institute of Psychiatry employees rather than professional actors. Both psychiatry residents and examiners found standardised patients acceptable, with both groups considering them competent to depict complex scenarios and accurate in their portrayal of psychiatric illness. The higher ratings provided by examiners in relation to competence may have been due to their involvement in writing case scenarios.

Table 3 Papers focussing on special topics in psychiatry OSCEs

Author year	Topic	Participants and OSCE characteristics	Key findings
Sauer et al. [21]	Attitudes of learners and teachers	• Psychiatry senior house officers (36) • Twelve 7-min stations	• Trainees evaluated OSCE positively, with 86% considering it fair and 89% appropriate in assessing clinical ability • Trainees preferred OSCE to the individual patient assessment it replaced in the RCPsych membership exam
Jabbar et al. [23]	Attitudes of learners and teachers	• Final-year medical students (111)	• OSCE was the most highly-rated teaching and examination method • The utility of OSCE in teaching and assessing clinical skills may have underpinned this preference
Sadeghi et al. [24]	Standardised patients	• Final-year psychiatry residents (21) and board-certified psychiatrist examiners (24) were surveyed • Eight 12-min stations	• Psychiatry residents and examiners found standardised patients acceptable, competent and accurate in their roles • Examiners (who wrote scenarios) rated competence more highly
Whelan et al. [25]	OSCE examiners	• Psychiatry trainees (55) • Eleven or 12 stations	• Moderate correlation between examiner and standardised patient scores for communication skills and overall performance • Examiners' scores were more strongly correlated • Including standardised patient scores in postgraduate psychiatry OSCEs required caution due to poor concurrent validity
Hodges and McNaughton [26]	OSCE examiners	• Commentary paper	• OSCE scoring by examiners involves a subjective, evaluative dimension • Emotionally-laden roles might affect standardised patients' perceptions of interviewers' competence • Research into suitability of OSCE examiners should incorporate broader educational and sociopolitical factors

(contined)

Table 3 (continued)

Author year	Topic	Participants and OSCE characteristics	Key findings
O'Connor et al. [28]	OSCE examiners	• Final year medical students (163) • Four five-min stations with simulated patients and two written stations	• Self-assessment of empathy was higher among female students • Higher concurrent validity between simulated patient (versus clinical examiner) assessment and self-assessment of empathy
Hanson et al. [29]	Child and adolescent psychiatry	• Medical students (402) • One of four 12-min child psychiatry stations in psychiatry clerkship OSCE	• Despite requiring considerable financial and staffing resources, substantial educational benefits were derived • Mean scores were 68–86% for content and 69–76% for process • Station reliability and examiner feedback were satisfactory
Hanson et al. [30]	Child and adolescent psychiatry	• Adolescent standardised patients in a medical student OSCE (39)	• Job skill acquisition and satisfaction of contributing to society identified as benefits • High stress psychosocial (substance abuse) role negatively received • No long-term adverse effects • Time invested in developing and pilot testing OSCE scenarios may reduce risks to adolescent standardised patients
Hung et al. [31]	Suicide risk assessment	• Psychiatry residents (26) and clinical psychology interns (5) • Single 40-min station	• A competency-assessment instrument for suicide risk-assessment (CAI-S) showed good internal consistency, reliability, and interrater reliability in an OSCE setting • Senior trainees performed better than junior trainees, supporting its concurrent validity
Matthews et al. [32]	Addiction psychiatry	• Third-year medical students • Two observed clinical interviews (OSCE format)	• An addiction psychiatry interclerkship training program improved knowledge, attitudes and confidence in this field • Skills in substance abuse assessment and intervention, as assessed by OSCE, were significantly improved

Padilla et al. [33]	Assessing cultural competence	• Psychiatry residents and fellows (17) • Observed clinical interview (OSCE format), one of two cases	• OSCE-based training familiarised trainees with culturally competent interviewing (centred on DSM-5 Cultural Formulation Interview) • Training incorporated a cultural formulation presentation and feedback
Moss [34]	Comprehensive psychiatric evaluation (single-station OSCE)	• Eleven first-year family medicine residents (five psychiatry training seminar participants) • Single observed clinical interview (OSCE format)	• Scenario revolved around a 47-year old separated woman with fatigue, abdominal pain and depressed mood • Superior OSCE performance (including diagnosis of major depression) among seminar participants
McLay et al. [35]	Comprehensive psychiatric evaluation (single-station OSCE)	• Third-year medical students (82) • Forty-five-minute interview, 15-min written test	• Performance on an OSCE simulating a detailed psychiatric interview correlated with students' essay examination, ward grades and scores on the National Board of Medical Examiners psychiatry examination • Students rated time availability, usefulness and believability of the process favourably
Chandra et al. [36]	OSCE as a teaching tool	• Thirty-four individual sessions over 6 months (15-min interview and 15-min group discussion and feedback)	• Objective structured clinical assessment with feedback (OSCAF) was convenient, economical and required few resources • Improved scoring, feedback and evaluation of effectiveness were recommended
Hodges [37] Hodges [3]	Contextual factors impacting on performance	• Theoretical papers	• Approaches to assessing OSCE validity may themselves be invalid • OSCEs identified as contextual, formative social experiences influenced by culture, economics and power relations • Qualitative research (cross-cultural approaches and analysis of sociological variables behind doctors' behaviours) needed to explain 'contextual fidelity'

5.3 Who Should Serve as Psychiatry OSCE Examiners?

In anticipation of the possibility that standardised patients might be called on to contribute to the scoring of postgraduate psychiatry OSCEs in the UK in future, Whelan et al. [25] measured the degree of agreement between scores given by examiners and standardised patients in two consecutive postgraduate mock OSCEs for 55 psychiatry trainees on a London psychiatry rotation. Standardised patients only allocated marks for communication skills and overall performance, whereas examiners also scored other skills and technical domains. Examiner and standardised patient scores for communication skills and overall performance were moderately correlated. Correlation between examiners' scores on these two domains was stronger. The authors concluded that including standardised patient scores in postgraduate psychiatry OSCE marking schemes required caution, as their moderate correlation with examiner scores was indicative of poor concurrent validity.

In a commentary on the paper by Whelan et al. [25], Hodges and McNaughton [26] argued that it is important to look beyond an OSCE's psychometric properties in considering who is most suited to the examiner role. A subjective, evaluative dimension is clearly present in OSCEs in which examiners complete global ratings or overall judgments of competence, and use of binary checklists to assess complex phenomena which exist on a continuum, such as empathy, rapport, and problem-solving, may be invalid and create an illusion of impartiality [27]. Furthermore, the process of scoring an OSCE performance is not entirely objective from a simulated patient perspective either. Some emotionally-laden roles (e.g. a patient with borderline personality disorder experiencing abandonment) might affect standardised patients' perceptions of interviewers' competence. Economic incentives were identified as a possible contributor to the popularity of standardised patient examiners in some settings. Incorporation of broader educational and sociopolitical factors into research regarding the appropriateness of OSCE examiners was recommended to ensure the fair appraisal of doctors' competence.

O'Connor et al. [28] compared 163 final-year medical students' self-assessment of empathy prior to their psychiatry OSCE at University College Dublin with assessments of empathy by clinical examiners and simulated patients during the OSCE itself. Self-assessment of empathy was significantly higher among female students. Concurrent validity was higher between simulated patient assessment (rather than clinical examiner assessment) and self-assessment of empathy, highlighting the potential validity of simulated patients as assessors of medical student empathy in an OSCE setting.

5.4 Child and Adolescent Psychiatry in the OSCE Setting

Hanson et al. [29] reported on the development and integration of four child psychiatry stations into a medical student psychiatry clerkship OSCE, designed to assess skills in recognising four common conditions. Child psychiatrists with

experience in OSCE prepared the case scenarios and marking checklists and supervised the training of standardised patients. Despite considerable financial and staffing resources being required, the educational benefits derived from the success of this initiative were substantial. Following the examination of 402 students, the mean scores were 68–86% for content and 69–76% for process, and station reliability and examiner feedback were satisfactory.

Recognising the potential for adolescent standardised patients to experience adverse simulation effects in psychiatry OSCEs, Hanson et al. [30] evaluated an adolescent standardised patient selection method and simulation effects resulting from low- and high-stress roles. A two-component standardised patient selection method excluded 21% of the 83 applicants (7% on employment and 14% on psychological grounds). Selected applicants were randomly assigned to a low-stress medical role (abdominal pain of infectious origin), a high-stress psychosocial role (substance abuse) or wait list control group. Thirty-nine standardised patients participated in a medical student OSCE at the University of Toronto. Acquisition of job skills and the satisfaction of contributing to society were identified as benefits of OSCE participation by standardised patients. Conversely, an unexpectedly strong negative perception of the substance abuse role emerged. No long-term adverse effects were noted. The authors concluded that time invested in developing and potentially pilot testing high-stress OSCE case scenarios may be worthwhile in reducing risks posed to adolescent standardised patients.

5.5 Suicide Risk Assessment in the OCSE Setting

Hung et al. [31] developed a competency-assessment instrument for suicide risk-assessment (CAI-S) and evaluated its use in an OSCE involving 31 trainees (26 psychiatry residents and five clinical psychology interns) at the University of California. The OSCE included a 15-min standardised patient interview; 15 min to write a progress note; a 10-min oral presentation, including an assessment summary and suicide risk management plan; CAI-S completion by the faculty assessor; and a 25-min discussion and feedback session for trainees. Good internal consistency, reliability, and interrater reliability were demonstrated for the CAI-S and the better performance of senior compared to junior trainees in the OSCE setting supported its concurrent validity.

5.6 Addiction Psychiatry in the OSCE Setting

In order address the problem of inadequate substance abuse teaching in the medical student curriculum, Matthews et al. [32] examined the immediate and delayed effects of an intensive 1 or 2-day substance abuse interclerkship program. The program integrated several teaching modalities (with a focus on small-group teaching)

and was delivered to third-year medical students at the University of Massachusetts Medical School between standard clerkship blocks. Improvements in knowledge, attitudes and confidence were demonstrated in pre- to post-interclerkship assessments. Skills in substance abuse assessment and intervention, as assessed by two OSCE interviews with simulated patients (one with and one without active substance abuse problems) at the end of a 6-week psychiatry clerkship, were significantly improved where the interclerkship had previously been completed.

5.7 Assessing Cultural Competence in Psychiatry OSCEs

OSCE-based training centred on the DSM-5 Cultural Formulation Interview was shown to be effective in familiarising 17 psychiatry trainees at the University of Massachusetts with culturally competent interview skills [33]. The training incorporated the presentation of a cultural formulation by trainees on which feedback was provided.

5.8 Comprehensive Psychiatric Evaluation in a Single-Station OSCE Format

In order to improve the teaching of psychiatric history taking, diagnosis, and management to first-year family medicine residents at the Sunnybrook Medical Centre (University of Toronto teaching hospital), Moss [34] developed and delivered a series of weekly, 1-h psychiatry seminars to six residents during their four-month Family Medicine Rotation. The program's effectiveness was evaluated via an OSCE conducted among 11 residents, five of whom had not participated in the seminar program. The single OSCE scenario revolved around a 47-year old separated woman with fatigue, abdominal pain and a depressed mood. Performance on history taking, communications skills and diagnostic and management skills was rated using a 100-point marking system that emphasised the importance of accurately diagnosing a major depressive episode. OSCE performance was significantly better among residents who participated in the seminar.

An OSCE designed to simulate a detailed psychiatric interview was used to assess 52 medical students at Tulane University School of Medicine [35]. Students were provided with the patient's chief complaint, a brief medical history and vital signs prior to undertaking a 45-min simulated patient interview. A written task encompassing differential diagnosis, safety risks and treatment planning was completed over a further 15 min. Students were graded using a 36-item content checklist and patient perception scale (both completed by the standardised patients) and a written examination component (completed by a psychiatrist). Checked independently by three fourth-year medical students who watched the interviews

on videotape (content checklist and patient perception scale) or an independent grader (written exam component), all three assessments correlated with students' essay examination and ward grades and scores on the National Board of Medical Examiners psychiatry examination. Students rated the time availability, usefulness and believability of the process favourably.

5.9 The OSCE as a Teaching Tool in Psychiatry

Chandra et al. [36] described the implementation of objective structured clinical assessment with feedback (OSCAF), an adaptation of OSCE for teaching purposes, at the National Institute of Mental Health and Neurosciences in Bangalore, India. OSCE adaptation involved language and cultural modification, use of supervised role play rather than standardised patients, and development of a 14-item checklist to guide feedback. While OSCAF was convenient, economical and required few resources, improvements in scoring and feedback, and evaluation of effectiveness, were recommended to establish its place in postgraduate psychiatry education.

5.10 Evaluation of Contextual Factors Impacting on Psychiatry OSCE Performance

In a theoretical paper regarding the validity of OSCE, Hodges [37] raised the possibility that approaches to assessing validity may themselves be invalid. OSCEs were identified as distinctly contextual and strongly formative social experiences that are significantly influenced by culture, economics and power relations. Sophisticated qualitative research, incorporating cross-cultural approaches and analysis of sociological variables underpinning doctors' behaviours, was suggested to be necessary to explain their 'contextual fidelity.' These themes were elaborated upon in a further paper by the same author [3].

6 Discussion

Review of the literature suggests that the OSCE has been widely adopted in psychiatry education in the English-speaking world. Available studies indicate that it can be a valid and reliable method of assessing competencies in psychiatry and is acceptable to both learners and teachers alike. Some shortcomings are readily apparent, however. Given its wide uptake in postgraduate (specialist) examinations, it is striking that much of the available research has been undertaken in undergraduate (medical student) psychiatry teaching. The validity of the OSCE

as a postgraduate psychiatry examination has been directly challenged [14]. For this reason, the recommendation that OSCEs be used in conjunction with other methods of clinical skills assessment [38] is particularly pertinent in the postgraduate setting. Problems in using OSCE as a form of summative assessment in residency training were identified since its inception [6]. While subsequent refinements in technique may have led to improvements in its validity as a summative assessment tool for medical students, it is unclear that this issue has been satisfactorily addressed in the postgraduate psychiatry domain. OSCE may be less prone to challenge when used exclusively as a tool for teaching [36] and providing formative assessment [6].

The aim of OSCE to examine practical competencies that are necessary for effective professional practice is undoubtedly well-intentioned. However, despite being labelled 'objective' and 'structured,' there is inherently greater subjectivity in the assessment of OSCE performance relative to other forms of assessment such a multiple choice questions, where absolute consensus regarding what constitutes a correct or incorrect answer is more easily achieved. In other words, highly objective but practically uninformative tests of knowledge have been replaced by the clinically relevant but potentially more subjective OSCE. Hodges [3, 37] identified a need to assess broader sociocultural and other contextual factors that may impact on performance in the psychiatry OSCE. Despite this call over a decade ago, qualitative research in this area is still lacking. Psychiatrists will be among the first to recognise the impact of 'unseen' psychodynamic factors [39] and sociocultural influences on human interactions outside the examination setting. The same variables are likely to come into play in the complex, high-stakes, emotionally-charged OSCE setting. However, subjecting these variables to scientific scrutiny may be far more complicated than merely identifying their potential existence.

With the above limitations in mind, the importance of adequately preparing both candidates and examiners for this form of assessment must be emphasised. OSCEs are very resource intensive to conduct and this may pose challenges in providing candidates and examiners with adequate practice for their respective roles. Furthermore, OSCE scenarios represent an idealised view of clinical practice and expecting candidates to acquire proficiency in performing them solely through observing and participating in busy clinical work schedules (as opposed to dedicated clinical skills teaching sessions) may be unrealistic. In the author's experience, groups of candidates will often establish their own OSCE preparation groups outside the formal teaching curriculum. Ideally, this should be complemented by formally instituted training sessions involving both candidates and examiners akin to those described by Robinson et al. [17].

It has been noted that creating high-quality OSCE stations requires time and effort [1]. It is important that OSCE stations are fit for the purpose for which they are designed. Relatively brief stations with checklist-like marking sheets may be suitable for examining discrete tasks in medical students (such as enquiring about manic symptoms) but may be inadequate for examining more sophisticated scenarios at a postgraduate training level (such as arranging assistance for an emotionally distressed junior colleague). Potentially longer stations and more sophisticated marking schemes are necessary for the latter. At the same time,

however, it is important from the candidates' perspective—and in the interests of fairness—that transparency about what is expected of them is maintained in presenting OSCE scenarios. A tendency to base stations on unusual or rare cases was one of the earliest identified challenges to the validity of psychiatry OSCEs [6] and remains pertinent to this day.

The benefits of 'reverse engineering' or 'deconstructing' OSCE stations became apparent to the author in the process of preparing for his own postgraduate psychiatry OSCE. This approach was based on the notion that a better understanding of how OSCE stations were developed increased the chances of exam success, and has recently been espoused in an entire textbook on the topic [40]. However, deconstruction may be ineffective if a station is poorly conceptualised and its required tasks not clearly presented. An understanding of the standard-setting method used for determining the OSCE pass mark may also be advantageous to both candidates and examiners [1].

OSCE may not be optimally suited to truly evaluating ethics and professionalism, as socially-expected responses may easily be learned and portrayed by candidates in simulated examination settings. The same limitation may be true, however, of other formal examination modalities. An essential, albeit subjective, adjunct to the appraisal of these domains may be supervisors' observation of students' and trainees' conduct in day-to-day clinical activities [41].

In light of the fatigue that traditional multi-station OSCE formats may cause in candidates and examiners [38, 42, 43], the question is raised of whether developing OSCEs with small numbers of in-depth stations would be advantageous. This approach might help to avoid sudden, repeated shifts in thinking, but its sampling of a smaller area of the curriculum may be a disadvantage. Two studies included in this review [34, 35] reported on the use of single-station OSCE formats involving lengthier patient interviews that were reminiscent of older-style clinical interview-based exams [44] but potentially better 'operationalised' in terms of their marking schedules. However, whether such single-station examinations should be regarded as OSCEs in the traditional sense—which imply a circuit of multiple stations—or classed in a category of their own remains open to debate, due to their sampling of a limited area of the curriculum.

Psychiatry, like medicine as a whole, has taken a leading role in adopting OSCE to assess clinical competencies. There is limited emerging evidence, however, of its adoption in the related field of psychology. Cramer et al. [45] recently proposed core competencies and an integrated training framework for suicide risk assessment training in doctoral psychology programs. At Monash University, with which the author is affiliated, both Clinical Psychology and Neuropsychology students are required to participate in two formative OSCEs and one summative OSCE during their Doctor of Psychology (DPsych) program, in keeping with recommendations for competency assessment by the Australian Psychology Accreditation Council (APAC) [46]. Sharma et al. [47] recently took the OSCE even further, describing their development of a Team OSCE (TOSCE) to promote interprofessional learning among psychiatrists, clinical psychologists and social workers in a mental health setting.

Despite any shortcomings, OSCEs are currently ubiquitous in all areas of undergraduate and postgraduate medicine and proposing a better alternative (other than returning to and refining former assessment methods) is difficult. The limited evidence base regarding the validity of OSCE in postgraduate psychiatry examinations suggests that more research is needed in this domain. A critical question is whether OSCE is sufficient on its own to assess high-level consultancy skills, and aspects of professionalism and ethical practice, that are essential for effective specialist practice, or whether it needs to be supplemented by additional testing modalities.

Funding No sources of funding were received for this work.

Conflicts of Interest The author reports no conflicts of interest and is alone responsible for the content and writing of this paper.

References

1. Boursicot, K.A.M., T.E. Roberts, and W.P. Burdick. 2014. Structured Assessments of Clinical Competence. In *Understanding Medical Education: Evidence, Theory and Practice*, ed. T. Swanwick, 2nd ed., 293–304. Chichester, West Sussex: Wiley.
2. Harden, R.M., and F.A. Gleeson. 1979. Assessment of Clinical Competence Using an Objective Structured Clinical Examination (OSCE). *Medical Education* 13: 39–54.
3. Hodges, B. 2003. OSCE! Variations on a Theme by Harden. *Medical Education* 37: 1134–1140.
4. ———. 2002. Creating, Monitoring, and Improving a Psychiatry OSCE: A Guide for Faculty. *Academic Psychiatry* 26: 134.
5. Davis, G.C. 2002. The Hodges Psychiatry OSCE Guide and Emerging Trends in Assessment. *Academic Psychiatry* 26: 184–186.
6. Loschen, E.L. 1993. Using the Objective Structured Clinical Examination in a Psychiatry Residency. *Academic Psychiatry* 17: 95–100.
7. Hodges, B., G. Regehr, M. Hanson, and N. McNaughton. 1997. An Objective Structured Clinical Examination for Evaluating Psychiatric Clinical Clerks. *Academic Medicine* 72: 715–721.
8. Hodges, B., and J. Lofchy. 1997. Evaluating Psychiatric Clinical Clerks with a Mini-Objective Structured Clinical Examination. *Academic Psychiatry* 21: 219–225.
9. Hodges, B., G. Regehr, M. Hanson, and N. McNaughton. 1998. Validation of an Objective Structured Clinical Examination in Psychiatry. *Academic Medicine* 73: 910–912.
10. Loschen, E.L. 2002. The OSCE Revisited: Use of Performance-Based Evaluation in Psychiatric Education. *Academic Psychiatry* 26: 202–204.
11. Barrows, H.S., R.G. Williams, and R.H. Moy. 1987. A Comprehensive Performance-Based Assessment of Fourth-Year Students' Clinical Skills. *Journal of Medical Education* 62: 805–809.
12. Park, R.S., J.T. Chibnall, R.J. Blaskiewicz, et al. 2004. Construct Validity of an Objective Structured Clinical Examination (OSCE) in Psychiatry: Associations with the Clinical Skills Examination and Other Indicators. *Academic Psychiatry* 28: 122–128.
13. Walters, K., D. Osborn, and P. Raven. 2005. The Development, Validity and Reliability of a Multimodality Objective Structured Clinical Examination in Psychiatry. *Medical Education* 39: 292–298. doi:10.1111/j.1365-2929.2005.02091.x.

14. Marwaha, S. 2011. Objective Structured Clinical Examinations (OSCEs), Psychiatry and the Clinical Assessment of Skills and Competencies (CASC) Same Evidence, Different Judgement. *BMC Psychiatry* 11: 1.
15. Blaskiewicz, R.J., R.S. Park, J.T. Chibnall, and J.K. Powell. 2004. The Influence of Testing Context and Clinical Rotation Order on Students' OSCE Performance. *Academic Medicine* 79: 597–601.
16. Park, R.S., J.T. Chibnall, and A. Morrow. 2005. Relationship of Rotation Timing to Pattern of Clerkship Performance in Psychiatry. *Academic Psychiatry* 29: 267–273.
17. Robinson, D.J. 2009. A Training and Examination Program to Prepare Psychiatric Residents for OSCE-Style Exams. *Academic Psychiatry* 33: 331.
18. Goisman, R.M., R.M. Levin, E. Krupat, et al. 2010. Psychiatric OSCE Performance of Students with and without a Previous Core Psychiatry Clerkship. *Academic Psychiatry* 34: 141–144.
19. Morreale, M., C. Arfken, P. Bridge, and R. Balon. 2012. Incorporating Active Learning Into a Psychiatry Clerkship: Does It Make a Difference? *Academic Psychiatry* 36: 223–225.
20. Griswold, T., C. Bullock, E. Gaufberg, et al. 2012. Psychiatry in the Harvard Medical School—Cambridge Integrated Clerkship: An Innovative, Year-Long Program. *Academic Psychiatry* 36: 380–387.
21. Sauer, J., B. Hodges, A. Santhouse, and N. Blackwood. 2005. The OSCE has Landed: One Small Step for British Psychiatry? *Academic Psychiatry* 29: 310–315.
22. Hodges, B., M. Hanson, M.N. McNaughton, and G. Regehr. 1999. What do Psychiatry Residents Think of an Objective Structured Clinical Examination? *Academic Psychiatry* 23: 198–204.
23. Jabbar, F., P. Casey, and B.D. Kelly. 2014. Undergraduate Psychiatry Students' Attitudes Towards Teaching Methods at an Irish University. *Irish Journal of Medical Science*: 1–4. doi:10.1007/s11845-014-1211-3.
24. Sadeghi, M., A. Taghva, G. Mirsepassi, and M. Hassanzadeh. 2007. How do Examiners and Examinees Think About Role-Playing of Standardized Patients in an OSCE Setting? *Academic Psychiatry* 31: 358–362.
25. Whelan, P., L. Church, and K. Kadry. 2009. Using Standardized Patients' Marks in Scoring Postgraduate Psychiatry OSCEs. *Academic Psychiatry* 33: 319–322.
26. Hodges, B.D., and N. McNaughton. 2009. Who Should Be an OSCE Examiner? *Academic Psychiatry* 33: 282–284.
27. Norman, G. 2005. Checklists vs. Ratings, the Illusion of Objectivity, the Demise of Skills and the Debasement of Evidence. *Advances in Health Sciences Education* 10: 1–3. doi:10.1007/s10459-005-4723-9.
28. O'Connor, K., R. King, K.M. Malone, and A. Guerandel. 2014. Clinical Examiners, Simulated Patients, and Student Self-Assessed Empathy in Medical Students During a Psychiatry Objective Structured Clinical Examination. *Academic Psychiatry* 38: 451–457. doi:10.1007/s40596-014-0133-8.
29. Hanson, M., B. Hodges, N. McNaughton, and G. Regehr. 1998. The Integration of Child Psychiatry Into a Psychiatry Clerkship OSCE. *Canadian Journal of Psychiatry* 43: 614–618.
30. Hanson, M., R. Tiberius, B. Hodges, et al. 2002. Adolescent Standardized Patients: Method of Selection and Assessment of Benefits and Risks. *Teaching and Learning in Medicine* 14: 104–113. doi:10.1207/S15328015TLM1402_07.
31. Hung, E.K., S.R. Fordwood, S.E. Hall, et al. 2012. A method for Evaluating Competency in Assessment and Management of Suicide Risk. *Academic Psychiatry* 36: 23–28.
32. Matthews, J., W. Kadish, S.V. Barrett, et al. 2002. The Impact of a Brief Interclerkship About Substance Abuse on Medical Students' Skills. *Academic Medicine* 77: 419–426.
33. Padilla, A., S. Benjamin, and R. Lewis-Fernandez. 2016. Assessing Cultural Psychiatry Milestones Through an Objective Structured Clinical Examination. *Academic Psychiatry* 40: 600–603. doi:10.1007/s40596-016-0544-9.
34. Moss, J.H. 1990. Evaluating a Seminar Designed to Improve Psychiatry Skills of Family Medicine Residents. *Academic Medicine* 65: 658–660.

35. McLay, R.N., P. Rodenhauser, M.D.S. Anderson, et al. 2002. Simulating a Full-Length Psychiatric Interview with a Complex Patient. *Academic Psychiatry* 26: 162–167.
36. Chandra, P.S., S.K. Chaturvedi, and G. Desai. 2009. Objective Standardized Clinical Assessment with Feedback: Adapting the Objective Structured Clinical Examination for Postgraduate Psychiatry Training in India. *Indian Journal of Medical Sciences* 63: 235. doi:10.4103/0019-5359.53391.
37. Hodges, B. 2003. Validity and the OSCE. *Medical Teacher* 25: 250–254. doi:10.1080/01421590310001002836.
38. Barman, A. 2005. Critiques on the Objective Structured Clinical Examination. *Annals Academy of Medicine Singapore* 34: 478–482.
39. Malan, D.H. 1979. *Individual Psychotherapy and the Science of Psychodynamics*. London: Butterworths.
40. Harding, D. 2014. *Deconstructing the OSCE*. Oxford: Oxford University Press.
41. Marrero, I., M. Bell, L.B. Dunn, and L.W. Roberts. 2013. Assessing Professionalism and Ethics Knowledge and Skills: Preferences of Psychiatry Residents. *Academic Psychiatry* 37: 392–397.
42. Rutala, P.J., D.B. Witzke, E.O. Leko, et al. 1990. Student Fatigue as a Variable Affecting Performance in an Objective Structured Clinical Examination. *Academic Medicine* 65: S53–S54.
43. Reznick, R., S. Smee, A. Rothman, et al. 1992. An Objective Structured Clinical Examination for the Licentiate: Report of the Pilot Project of the Medical Council of Canada. *Academic Medicine* 67: 487–494.
44. Flynn, B. 2015. The Abandonment of the RANZCP OCI – and Why We May Live to Regret It. *Australasian Psychiatry* 23: 303–305. doi:10.1177/1039856215581295.
45. Cramer, R.J., S.M. Johnson, J. McLaughlin, et al. 2013. Suicide Risk Assessment Training for Psychology Doctoral Programs: Core Competencies and a Framework for Training. *Training and Education in Professional Psychology* 7: 1–11. doi:10.1037/a0031836.
46. Monash University. 2016. Psych: Graduate Research Info Guide – 10. DPsych OSCE. http://www.med.monash.edu.au/psych/students/current/hdr-info-guide/dpsych-osce.html. Accessed 8 Aug 2016
47. Sharma, M.K., P.S. Chandra, and S.K. Chaturvedi. 2015. Team OSCE: A Teaching Modality for Promotion of Multidisciplinary Work in Mental Health Settings. *Indian Journal of Psychological Medicine* 37: 327. doi:10.4103/0253-7176.162954.