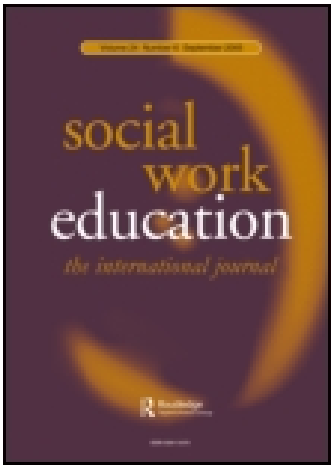


Social Work Education: The International Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cswe20

# A New Methodology for Assessing Social Work Practice: The Adaptation of the Objective Structured Clinical Evaluation (SW-OSCE)

Yuhwa Eva Lu [a], Eileen Ain [b], Charissa Chamorro [a], Chiung-Yun Chang [a], Joyce Yen Feng [c], Rowena Fong [d], Betty Garcia [e], Robert Leibson Hawkins [a] & Muriel Yu [f]

[a] New York University, USA
[b] Private Practice, New York City, USA
[c] National Taiwan University, Taiwan
[d] University of Texas at Austin, USA
[e] California State University, Fresno, USA
[f] University of Texas at Arlington, USA

Published online: 06 Feb 2011.

To cite this article: Yuhwa Eva Lu, Eileen Ain, Charissa Chamorro, Chiung-Yun Chang, Joyce Yen Feng, Rowena Fong, Betty Garcia, Robert Leibson Hawkins & Muriel Yu (2011) A New Methodology for Assessing Social Work Practice: The Adaptation of the Objective Structured Clinical Evaluation (SW-OSCE), Social Work Education: The International Journal, 30:02, 170-185, DOI: 10.1080/02615479.2011.540385

To link to this article: http://dx.doi.org/10.1080/02615479.2011.540385

*Social Work Education*
*Vol. 30, No. 2, March 2011, pp. 170–185*



# A New Methodology for Assessing Social Work Practice: The Adaptation of the Objective Structured Clinical Evaluation (SW-OSCE)

Yuhwa Eva Lu, Eileen Ain, Charissa Chamorro, Chiung-Yun Chang, Joyce Yen Feng, Rowena Fong, Betty Garcia, Robert Leibson Hawkins & Muriel Yu



*The Objective Structured Clinical Evaluation (OSCE) methodology was originally developed to assess medical students. OSCE is a carefully scripted, standardized, simulated interview, in which students' interactional skills are observed and assessed. Here it is examined for its potential use in assessing social work practice skills. The development of the Social Work OSCE (SW-OSCE) and the Clinical Competence-based Behavioural Checklist (CCBC) are described. Findings from a pilot study assessing MSW students' clinical skills with explicit observable criteria of the CCBC are presented.*

*A quantitative and qualitative mixed-methods data analysis was applied. The CCBC had high internal reliability, for both the overall sample and for the different case scenarios, with Cronbach's alpha values ranging from 0.888 to 0.965. The validity of the instrument was also examined: qualitative content analysis of the taped interviews indicated that clinical skills and cultural empathy are not synonymous. The racial/ethnic match between the student and the 'client' did not predict better rapport or more cultural empathy. Examination grades are not necessarily consistent with actual performance in either clinical competence or cultural empathy or vice versa.*

*Nevertheless, the results provide some support for the use of the SW-OSCE as a tool for assessing performance in social work practice. They also indicate its potential for evaluating the outcomes of educational programmes.*

Yuhwa Eva Lu, Charissa Chamorro, Chiung-Yun Chang & Robert Leibson Hawkins, New York University, USA; Eileen Ain, Private Practice, New York City, USA; Joyce Yen Feng, National Taiwan University, Taiwan; Rowena Fong, University of Texas at Austin, USA; Betty Garcia, California State University, Fresno, USA; Muriel Yu, University of Texas at Arlington, USA

Correspondence to: Yuhwa Eva Lu, Associate Professor, Silver School of Social Work, New York University, 1 Washington Square North, New York, NY, 10003, USA. Tel.: 212-998-9094. Fax: 212-995-4588; Email: eva.lu@nyu.edu

ISSN 0261-5479 print/1470-1227 online 
DOI: 10.1080/02615479.2011.540385

*Keywords: Objective Structured Clinical Evaluation; Assessment of Social Work Students; Cultural Empathy; Outcomes of Social Work Education*



## Introduction

Social work educators have long sought valid and reliable ways to evaluate direct practice, particularly for working with vulnerable populations and programme evaluation (i.e. outcomes) in social work education. Two social work educational outcomes that are exceptionally difficult to assess in ways that capture the related complexity and nuances are students' practice competence and their cultural empathy, particularly in real time interactions (Cusimano *et al.*, 1998; Gambrill, 2001; Bogo *et al.*, 2002; Holden *et al.*, 2002; Carpenter, 2005).

Several decades ago, Elstein and colleagues (1978) reported that knowledge of a particular medical condition was more directly related to a physician's performance than to their general problem-solving ability. This work set the stage for the subsequent 20-year dominance of the Empirical-Research Model in US social work schools' efforts to evaluate practice (Ventimiglia *et al.*, 2000). Still, this model is apt to neglect the more subtle interactional, dynamic and contextual elements of the client–social worker relationship process. Although practice evaluations bridge the gap between research and practice, findings indicate that discrepancies exist between student self-report and observational assessment, self and peer assessment, and student and instructor assessment (Ellis, 2001; Herie and Martin, 2002; Regehr *et al.*, 2002; Tousignant and DesMarchais, 2002).

While grades in direct social work practice courses and the field internship experience provide the means to assess practice performance, instructors often rely on paper and pencil tests completed by the student. While the mid-term and final evaluations may be effective for assessing the domains of knowledge and cognitive reasoning, these measures have become problematic in the US due to faculty subjectivity and grade inflation (Shoemaker and DeVos, 1999; Noble and Stretch, 2002; Scanlan and Care, 2004). Thus, the lack of a systematic methodology in teaching practice skills and for the evaluation of practice competency makes it difficult to identify students who are unsuitable for the social work profession, as well as to evaluate the effectiveness of teaching courses and programmes (O'Hare and Collins, 1997; Ryan *et al.*, 1997; O'Hare *et al.*, 1998; Lafrance *et al.*, 2004).

The issue of early identification of students who are a poor match for their identified profession is not unique to social work. The field of medical education has addressed the issue with the use of Objective Structured Clinical Evaluation (OSCE) (Harden and Gleeson, 1979). The OSCE is a carefully scripted, standardized, simulated interview in which students' interactional skills are observed and assessed. These interview sessions are video-taped and followed by a post-session debriefing. OSCE has three fundamental elements: (1) a simulated patient played by an actor; (2) direct observation of performance; and (3) assessments of their performance using a

competence-based behavioural checklist. The checklist indicators guide instructors on the specific behaviours to evaluate.

The OSCE applies Social Cognitive Theory, which describes learning in terms of the relationship between behavioural, environmental and personal factors (Bandura, 1989). According to this theory, interactive learning gives students opportunities to increase self-efficacy through practice (Leopold *et al.*, 2005) making it a method for learning as well as an assessment tool. OSCE has been used for over a decade as a gateway assessment tool for licensing exams in medicine in Australia, Canada and the United States (Reznick *et al.*, 1998).

Earlier evaluations of the use of OSCE have shown that it helps to enhance student confidence as well as competence. Furthermore, OSCE pre-test and post-test results have been successful in predicting learning processes and educational outcomes (Ytterberg *et al.*, 1998). According to Hawker *et al.*, (2010), OSCEs could be reliable predictors for assessing learning outcomes. In their six-year study of dietetic students in England they reported that the preclinical OSCE was a significant predictor of students' subsequent performance; over 90% of students reported positive feedback from the OSCE experiences on their clinical skills.

Ogawa *et al.*, (2003) reported that the OSCE shows high face validity, content validity and reliability. However, a critical review by Turner and Dankowski (2008) subsequently concluded that the OSCE has wide variation in reliability scores and that its validity is inconclusive. Until now, no standardized checklist for a social work OSCE has been reported in the literature, but it is clear from the above that any such work must be accompanied by careful attention to reliability and validity. In this paper, we apply the OSCE approach to social work. The specific objectives are: (1) to describe the development of the Social Work Objective Structured Clinical Evaluation (SW-OSCE), an adaptation of the OSCE for the social work profession; and (2) to report the preliminary findings from a pilot study.



## Adaptation of the OSCE Methodology to Social Work

The social work adaptation of OSCE is built on the concept of evidence-informed and competency-based assessment. The OSCE methodology moves beyond didactic teaching and into actual practice and development of core skills, thus transforming intellectual learning into new behaviours. Similarly, the focus of the SW-OSCE is to establish reliable and concrete criteria for assessing students' actual performance. The SW-OSCE uses a social work competence-based behavioural checklist comprised of the constellation of knowledge, values and skills considered to define social work education. SW-OSCE, the assessment of performance in the brief interview, has focused on measuring social work core skills, however, it is difficult to reach a consensus of what these practice skills are in social work practice.

American researchers Koroloff and Rhyne (1989) developed a 25-item social work competence-based behavioural checklist, which comprised four subscales: interpersonal communication, assessment, intervention and termination skills. Subsequently, O'Hare and colleagues (1998) developed a Practice Skills Inventory (PSI) to assess

how frequently social workers use supportive, therapeutic and case management skills. Bogo and colleagues (2004) conducted a Practice-Based Evaluation (PBE) which categorizes six dimensions of a social work practice competence tool. These cover learning and growth, behaviour in the organization, clinical relationships, conceptualizing practice, assessment and intervention, and professional communication.

Similar efforts in the social work profession have aimed at defining 'cultural competence' in addition to a focus on practice skills assessment. Included in these cultural competency assessments are the *Indicators for the Achievement of the NASW Standards for Cultural Competence in Social Work Practice* (2007) developed by the National Association of Social Workers (NASW) National Committee on Racial and Ethnic Diversity (NCORED). The Council on Social Work Education (CSWE) has recommended that the NASW Indicators, and the new *Educational Policy and Accreditation Standards* (2008) include a competency on addressing diversity and difference in practice skills.

As in medicine, 'simulated clients' have been utilized in social work education. Badger and MacNeil (2002) developed this approach in order to allow social work students to practise assessment skills for interviewing clients with mental health problems. They considered that such a simulated assessment system provides an opportunity to control the characteristics of case interactions, to observe student–client interaction over a broad range of client types and situations, and, overall, to assess student performance. Similarly, Miller (2004) developed two simulated client scenarios to assess assessment and referral skills. Baez (2005) reported the application of OSCE to social work education with a focus on substance abuse and teaching students screening and intervention skills. Students rated the overall experience of participation in the OSCE extremely favourably, with a mean of 4.9 on a five-point scale. However, so far there are no reports regarding any standardized instrument developed for social work OSCE applications. Currently, as the social work profession evolves to the standards of evidence-informed and competence-based practice, researcher approaches include set criteria, measurable behavioural indicators and prescribed learning outcomes. Therefore, the authors of this study endeavoured to develop a Clinical Competence-based Behavioral Checklist (CCBC) for social work practice.



## Methodology

### *Instrument Development*

There were several phases in adapting the OSCE methodology.

#### *Clinical Competence-based Behavioral Checklist (CCBC)*

Before SW-OSCE could be implemented and fully employed, a 'competence-based behavioral checklist' (CCBC) had to be established based on theories, concepts and empirical data. The first step was to develop categories that reflected the values of practice competence and cultural sensitivity in the social work profession. A literature review, survey and focus group yielded 10 categories of professional competence

criteria: professional values, knowledge, cultural empathy, interviewing skills, intervention skills, empowerment perspectives, critical thinking, professional use of self, evaluations and knowledge of legal mandates (Medina *et al.*, 2004).

Some researchers have reported that overemphasizing objectification and creating a long list of criteria to be tested trivializes the assessment and decreases the validity of the OSCE; overly detailed checklists only give the appearance of objectivity by measuring the thoroughness of the task rather than the performance (Wilkinson *et al.*, 2003). Thus, to enhance the operational efficiency of the measurement instrument, the 10 categories identified by our team (Medina *et al.*, 2004) were consolidated into four: (1) interviewing skills; (2) cultural empathy; (3) assessment and intervention strategies; and (4) comprehensive evaluation. To better assess the accuracy of students' insights and critical thinking about their own performance, a 'metacognition' category was added. Metacognition is defined as the 'general ability to recognize one's own performance levels of competency and bring learning processes into consciousness' (Nisbet and Shucksmith, 1984). For the purpose of this study, we are defining metacognition as the ability to recognize the strengths and weaknesses of one's own performance and level of clinical competence. It follows that competence-based social work practitioners are positive, productive, able to self assess and aware of their thought processes (Maudsley and Strivens, 2000).

To increase the reliability of the measure, an 'overall score' category was added. This assesses a broader set of skills than individual categories. Regehr *et al.*, (1999) have reported that the scores for an overall assessment category are at least as reliable as the scores for individual categories and in some cases more valid. Finally, a qualitative 'comments section' was also added for assessors to record their scoring rationale (see Chart 1).

The CCBC takes between three and five minutes to complete and evaluates student performance on a nine-point, Likert scale with one (1) reflecting a 'poor performance', and nine (9) reflecting an 'excellent performance'. No item is reverse coded and total scores for the five categories can range from 5 to 45.

For each category, a list of behavioural indicators guides the rater. Thus, the 'interviewing skills category' measures the students' use of reflective listening, verbal and non-verbal communication, and professional use of self. To receive a high score in this category, students exhibit good reflective listening skills, focus on gathering factual information, are expressive without an assumptive attitude and do not give advice to the simulated client. The 'cultural empathy category' focuses on the students' ability to demonstrate awareness of their own culture, sensitivity to diverse cultural values and constructs, and proficiency in cross-cultural communication. Students who receive high scores in this category are able to display awareness of power differentials and oppression, elicit and affirm the client's cultural identity, and then formulate culturally congruent assessment and intervention strategies.

The 'assessment and intervention category' is used to evaluate whether the student can effectively assess the client's needs, develop goals, lay out steps toward those goals, and use multiple intervention models. A successful student interviewer will be aware of and 'address' the client's immediate needs, clarify practical goals with the client, lay out steps

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Interviewing Skills**<br>• Reflective Listening<br>• Verbal and non-verbal communication<br>• Professional Use of Self<br>• Explaining the purpose of the meeting | Poor listening skills; hindering of information gathering; assumptive attitude; giving of advice | | | Moderate level listening; attempt to gather information; expression of sympathy | | | Reflective listening; focus on factual information gathering; expression of empathy | | | Comments For Scoring |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| **Cultural Competence**<br>• Self awareness of clinician's own culture<br>• Sensitivity to diverse cultural values and constructs<br>• Proficiency in cross cultural communication | Insensitive to client's cultural values; judgmental attitude; imposition of ideas | | | Recognition of cultural differences; awareness of power differentials and oppression; ability to discern client's values | | | Respect for client's cultural values; ability to elicit and affirm cultural identity; professional judgment formulation | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| **Knowledge and Intervention Strategies**<br>• Assessment of client's needs/challenges and strengths/social resources<br>• Develop mutually agreed goals<br>• Set priorities and small steps for goal attainment and assess level of severity<br>• Demonstrate use of different models | Unaware of client's needs; lack of self direction; setting of impracticable goals | | | Awareness of client's expressed needs; limited use of practice models; attempt to develop goals for client | | | Ability to address client's immediate needs; reframing and empowering client through eclectic practice models; clarification of goals, and small action steps for goal attainment | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| **Evaluation**<br>• Identify indicators of client's total quality of life (i.e., health, mental health and economic status)<br>• Timely closure, termination of Referral | Failure to assess client's quality of life; inability to monitor progression of change; failure to identify indicators for termination or to make a referral | | | Awareness of client's prognosis and the indicators for quality of life; postponement of closure, termination or referral | | | Acknowledgement of client's progression of change; ability to make timely closure, termination and referral | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| **Meta Competence**<br>• Ability to recognize one's own performance and level of clinical competence | Inflated assessment of, or underestimation of, one's own performance and clinical competence | | | Appropriate assessment of one's own clinical performance and competence | | | Identify specific processes; recognize the dynamics of interaction with the client; aware of how to improve | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| **Overall Rating** | | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |

**Chart 1** CCBC—Checklist for SW-OSCE.

toward attaining those goals, and reframe and empower the client through multiple practice models.

The 'comprehensive evaluation category' allows the rater to assess the student's ability to identify indicators of the client's overall quality of life. To receive a high score, the student must acknowledge the client's progression of change, make a timely closure and make appropriate referrals.

The last category, 'metacognition' measures the ability to recognize one's own performance and level of clinical competence. It is scored after the student's brief post-interview discussion with the course instructor. Low-scoring students either inflate or underestimate their performance and competence, while successful students appropriately assess their performance and competence, identify specific processes, recognize the dynamics of their interaction with the client, and demonstrate awareness of how to improve.

*Case-scenario development*

Six scenarios were developed to evaluate a MSW course in 'social work practice with diverse populations' at New York University in Fall 2005. Each case scenario reflected diversity in clients' age, gender, race, religion, sexual orientation and social economic status (SES). Each comprised a 10–15 page detailed transcript of the client's emotional

state of being and life story. In the first scenario, 'Ms Stein', a 45-year-old Orthodox Jewish woman, became anxious, angry, fearful and panicked after being told that her adolescent daughter was pregnant. In the second, 'Ms Lee', a 35-year-old immigrant from China, learned of her daughter's truancy and poor academic performance. In the third scenario, 'Mr Shayan', a 20-year-old Iranian international student, felt hopeless, helpless and sleepless, and had substance abuse and motivational problems. In the fourth scenario, 'Ms Williams', a 34-year-old African American woman of Caribbean descent, felt anguished and guilty and had trouble maintaining relationships with her fiancé and family after being diagnosed as HIV positive. In the fifth scenario, 'Ms Perez', a 59-year-old American-born Latina from high SES background, reported that she was experiencing depression and fear regarding her future. In the sixth scenario, 'Mr Rodriguez', a married 24-year-old American-born Puerto Rican, felt confused about his sexual orientation and was mistrustful of a social worker.



### *Procedures for Administering SW-OSCE*

#### *Preparation phase*

The team developed detailed scripts of case scenarios, and a one-page brief instruction sheet with each client's background and presenting problem. We then hired actors as simulated clients, trained them in the procedure, and familiarized them with the definition of each category, and the 52 indicators and the scoring system of the CCBC. The purpose was that the actors would be able to rate the students. One month prior to the interviews, the actors were given the case-scenarios and 'pre-planned prompts' that contained specific words to be used verbatim and key information to convey to the students during the interview. Each actor participated in a practice session to ensure consistent interpretation of the interview scenario. Video-taping was arranged ahead of time and if possible in a room with a one-way mirror to emulate a confidential interview and to minimize distractions. One week before the interviews, the CCBC was discussed by the students and the class was prepared for the protocol.

#### *Protocol*

On the day of the SW-OSCE, student interviewers from the same class waited outside the classroom. They were prompted one by one to enter the classroom and remained there following their interview. The remainder of the class observed behind a one-way mirror (if available) or in the room. By plan, the interview scenario selected for each cohort of student interviewers was new and unfamiliar; it was often based on the actors' availability. After the student interviewers had signed the written consent forms, the instructor distributed the one-page information sheet, inclusive of the client's background and presenting problems, to each student interviewer. Each student was then ushered in for the 10–12 minute interview. A warning bell was sounded two minutes prior to the end of the interview to indicate that students should wind it up. For between three and five minutes immediately following the interview, the instructor asked probing questions about the student's self-evaluation of the

interview. Through this procedure, the student was able to self-assess their level of engagement with the client, cultural sensitivity, possible assessment and intervention strategy, level of satisfaction with the performance, and what he or she would do differently if given another opportunity. Between four and five interviews were conducted during each 100-minute class session. Four groups of evaluators—the instructor, the simulated (actor) client, the student peers and the student interviewer—rated the student's performance following each interview.

The final part of the protocol was a debriefing which was designed to maximize the students' learning experience. During this half hour period, all parties, including the instructor, students, actors and peers, talked about their experience with the interview. During this period, insights, receptivity levels to comments and experiential learning were paramount. Participating students were then given a written assignment of approximately six pages in which they were required to reflect on the OSCE exercise.



### *Participants*

All 101 student interviewers were MSW student volunteers, choosing SW-OSCE as the alternative final assignment for a required course, 'Ethno-cultural Issues in Social Work Practice'. All the students who took this course between 2006 and 2010 had the option of participating in the SW-OSCE. Their ages ranged from 31 to 56, with a mean of 31 and a median of 28-years-old. Most students were in their first semester, two-thirds had no previous social work experience, while one-third had some experience—from 3 to 16 years.

A total of 316 CCBC ratings were available for analysis. These were self-ratings made by the students ($N = 101$), the actors ($N = 109$) and the instructors ($N = 106$). Because there were some missing data in the student interviewer category, we have included only 109 partially matched CCBC ratings by students, actors and instructors. For this pilot study, the qualitative data analysis was based on a comparison of (20) students' interviews of the Ms Williams' case-scenario, performed by the same actor:

> Ms Williams, a 34-year-old, religious, immigrant African American woman of Caribbean descent, felt anguished, was confused, and had trouble sleeping and maintaining a normal level of daily functioning after being diagnosed as HIV positive.

### *A Mixed-Method Data Analysis*

The overall study applied a combination method of qualitative and quantitative data analysis. Statistical analysis of ANOVA was used for CCBC development. The qualitative content analysis of this study involves: (1) examining the specification of the content categories and themes from the transcripts; (2) applying explicit steps to compare these themes and categories with the CCBC; and (3) using the guided interview method to score, summarize and deliberate on the taped interviews with the research team.

*Quantitative Data Analysis*

*Reliability*

The internal consistency of the CCBC was evaluated using Cronbach's alpha coefficient with a 95% confidence interval. The alpha value was 0.946 for the overall rating categories, and 0.888–0.965 for the other categories which were interviewing skills, cultural empathy, assessment and intervention strategies, comprehensive evaluation and closure, and metacognition. Inter-rater reliability was also assessed and was reported in an earlier paper (Lu *et al.*, 2006); research inter-rater consistency ranged from 91% (overall score category) to 73% (metacognition category).

To determine the possible data-merge and to understand whether or not the six scenarios were equally challenging for the students, a one-way analysis of variance (ANOVA) was applied. There were statistically significant differences in their mean scores on two of the five categories: 'cultural empathy' and 'assessment-and-intervention strategies' (see Table 1). Considering these statistically significant differences, the cross-case scenario comparison was not conducted.

The mean scores of the students' self-ratings were compared with the ratings of the simulated clients and the instructors using ANOVA (Table 2). In general, students scored themselves significantly higher on both 'cultural empathy' and 'metacognition' categories than the actors and the instructors, in that order, except for the Asian-American case scenario. In this context it may be relevant that the instructor of the course is a bi-lingual, bi-cultural Asian-American woman.

In the 'metacognition category' the data showed about 36% respondent consistency rates. The students with high performance scores tend to rate themselves slightly lower than other raters, while students with lower mean scores regularly rated themselves higher than the other raters on all categories. This discrepancy raises concern regarding suitability for the helping professions and deserves further investigation. We consider all students with consistent scoring with other observers as having good 'metacognition', even the ones who perform poorly; we believe that these students are likely to improve because they recognize their own weaknesses. This perspective highlights the value of self-awareness and willingness to reflect as the basis of on-going and future skill learning.

*Validity*

It is suggested that a minimum of 150 subjects is required to perform confirmatory factor analysis of a scale (Cliff, 1987). Although the OSCE method often utilizes several interviews, our study only utilized one interview session per student resulting in a small number of sessions. Due to the small sample size, the validity of the CCBC was reviewed through a qualitative content analysis. The findings are reported in the following section.

*Qualitative Content Analysis*

The research team applied a two-stage content analysis method. First, each rater watched the video-tape recordings individually and wrote comments. Then, based on verbatim transcripts, all raters coded the texts into themes, subthemes and categories,

**Table 1** CCBC—ANOVA for Five Scenarios ($N = 316$).

| | Stein ($N = 113$) | | Lee ($N = 71$) | | Shayan ($N = 41$) | | Williams ($N = 50$) | | Perez ($N = 41$) | | Total ($N = 316$) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *F*(4,311) |
| 1. Interviewing skills | 6.81 | 1.562 | 6.51 | 1.620 | 6.50 | 2.214 | 7.09 | 1.603 | 6.17 | 1.687 | 6.66 | 1.733 | 2.136 |
| 2. Cultural empathy | 6.50 | 1.571 | 5.98 | 1.661 | 5.49 | 2.026 | 5.52 | 1.961 | 4.96 | 2.014 | 5.97 | 1.883 | 7.161*** |
| 3. Assessment & intervention strategies | 6.41 | 1.730 | 6.33 | 1.705 | 6.10 | 2.417 | 6.68 | 1.812 | 5.35 | 1.776 | 6.29 | 1.893 | 3.405* |
| 4. Comprehensive evaluation and closure | 6.49 | 1.764 | 6.38 | 1.737 | 6.16 | 2.215 | 6.30 | 1.657 | 5.62 | 1.900 | 6.31 | 1.857 | 1.788 |
| 5. Metacognition | 7.24 | 1.412 | 6.72 | 1.487 | 6.38 | 2.058 | 6.84 | 1.621 | 6.51 | 1.780 | 6.82 | 1.703 | 3.110* |
| 6. Overall evaluation | 6.96 | 1.439 | 6.53 | 1.602 | 6.28 | 2.168 | 6.79 | 1.466 | 6.09 | 1.583 | 6.65 | 1.666 | 3.014* |
| Cronbach's alpha of the checklist | 0.938 | | 0.952 | | 0.965 | | 0.888 | | 0.926 | | 0.946 | | |

*Notes*: ANOVA with Scheffe. Post hoc multiple comparisons were used to test the differences among means. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

**Table 2** CCBC—Comparison of Three Groups of Raters ($N = 316$).

| | Client rater ($N = 106$) | | Faculty rater ($N = 109$) | | Student self-rating ($N = 101$) | | Total ($N = 316$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *F*(2,313) |
| 1. Interviewing skills | 6.64 | 1.740 | 6.72 | 1.909 | 6.63 | 1.433 | 6.67 | 1.707 | 0.087 |
| 2. Cultural empathy | 6.05 | 1.904 | 5.54 | 1.992 | 6.12 | 1.564 | 5.90 | 1.847 | 3.141* |
| 3. Assessment & intervention strategies | 6.45 | 2.001 | 6.25 | 2.042 | 6.06 | 1.508 | 6.26 | 1.874 | 1.137 |
| 4. Comprehensive evaluation and closure | 6.39 | 1.802 | 6.25 | 2.065 | 6.20 | 1.598 | 6.28 | 1.833 | 0.296 |
| 5. Metacognition | 6.91 | 1.432 | 6.56 | 1.990 | 7.12 | 1.321 | 6.85 | 1.628 | 3.284* |
| 6. Overall evaluation | 6.79 | 1.643 | 6.52 | 1.904 | 6.59 | 1.254 | 6.63 | 1.629 | 0.793 |

*Note*: ANOVA with Scheffe. Post hoc multiple comparisons were used to test the differences among means. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

using the open coding guidelines proposed by Strauss (1987). In the second stage, a guided-content-analysis method was applied. Each rater scored the taped interviews using the CCBC. Subsequently and as a group, all raters reviewed the taped interviews. The researchers arrived at a consensus on a list of codes and matched these themes and subthemes with the 52 indicators of the CCBC checklist. The authors were able to find matches for all themes except three items: (1) clarification of the interviewer's function and the helping process; (2) application of a strength-based and focused attention on client's assets and resources; and (3) using an accurate scaling system to assess the duration and intensity of the client's presenting problem. The CCBC checklist has since been revised to include the new criteria (see Chart 1).



### *Additional Findings from the Qualitative Analysis*

After much deliberation, the following consensus was reached. First, clinical skills and cultural empathy are not synonymous; a high overall CCBC score did not assure higher perceived competence in cultural empathy. Second, a racial/ethnic match between the student and the client did not predict greater cultural empathy or better rapport. Third, incoming students without training may perform with competence, both clinically and culturally, so prior experience or professional training in social work are *not* indispensable for either practice competence or cultural empathy, and test grades were not necessarily consistent with actual performance in either practice competence or cultural empathy, or vice versa.

In each semester, the client/actor was asked to choose the best among the (four or five) interviewers; they consistently preferred the one with the highest score for cultural empathy. It appears that in the initial phase of the interview, development of the working alliance relies on cultural empathy more than other categories of practice skills.

Another important finding was that most students demonstrated varying levels of clinical competency (such as, reflective listening, physical attentiveness, empathic engagement, application of a strengths-based approach, multifaceted assessments and basic awareness of the clinical process). This reflected the core curriculum content of the social work practice courses of the MSW programme in the US. However, even with the presenting problems and life history of the client Ms Williams, most students did not explore how the client contracted HIV (i.e. chemically dependent, sexual history, etc.), or her immigrant status, and the withdrawal from her major social support from her faith community. Although sometimes addressed in different individual classes, none of these content areas are part of the social work core curriculum.

## Discussion and Implications

Although more development and analysis is needed, the findings of this pilot study indicate that the CCBC has potential as an objective, structured and competence-based

checklist for clinical assessment. The SW-OSCE is a viable supplement to traditional paper and pencil testing methods.

However, there are some challenges in the application of the SW-OSCE. A SW-OSCE session is relatively labour-intensive to create, and costly to administer. It could be more cost-efficient to develop the tool for a larger number of users and for many candidates to be examined during one administration. In our study, a non-funded research project, we used only a single-scenario design, which does not adequately assess five categories of the CCBC. Additionally, when a rater is also one of the course instructors, the ratings may be skewed. Each rater brings to the student's assessment his or her own cultural and personal bias. Raters may or may not be aware of these biases and need to separate their biases from student performances. An earlier pilot study of SW-OSCE (Lu *et al.*, 2006) indicated that the assessment of the student's cultural competence is based on both the student's actual performance and the rater's biases or expectations. Ideally, to minimize rater bias, there should be multiple raters assessing students from classes other than their own. An important next step is to use faculty members from diverse backgrounds (i.e. race, ethnicity, gender, age and other diversity).

To date two scenarios of the SW-OSCE and the CCBC form have been translated into Hebrew and Chinese, and similar versions of the study were done in Israel, Taiwan and California, USA. Yet, the replicability of the SW-OSCE still must be demonstrated. There is a need for consistency in administering the SW-OSCE, as well as ensuring inter-rater reliability. However, further work is required on assessing the internal reliability of SW-OSCE and investigating its construct validity through a principle components analysis of a large data set.

SW-OSCE, like all psychometric instruments, is not a perfect tool. As Turner and Dankowski (2008) pointed out, it is difficult to make conclusive statements about the validity of the OSCE method, since many variables in the design and implementation will influence the validity. The scoring methods for OSCE, which varied widely, will influence reliability; and no matter what method of scoring is used, there is always a concern regarding inter-rater reliability. To ensure high validity and reliability, one must attend to a large number of scenarios and raters, design specific scenarios for each category of skills, with well-trained patients/clients and well-designed implementation procedures. The future development of the SW-OSCE scenarios needs to go beyond the medical-individual model and expand to the couple, family and small group contexts, as well as to focus on diverse arenas of social work practice, such as, different age populations, presenting problems, cultural diversity training, crisis intervention/safety training, and various speciality fields of concentration.

Finally, with particular reference to the focus of this special issue, we can note the potential use of the SW-OSCE in evaluating the outcomes of social work education courses. Lozano *et al.* (2010) recently reported a small scale ($n = 18$) pilot study in the US which provides a possible model. They evaluated a brief training course (two sessions of 4.5 hours) in motivational interviewing (MI) for junior paediatricians using a randomized controlled trial with a waiting list control group. Three OSCE scenarios were developed and employed to assess participants at the baseline, three

months and seven months later. The researchers reported behaviour change in the form of improved counselling following the MI training and personal feedback to the participants. The indications are that this grant-funded study should be a pilot for a large scale trial: the methodology is sophisticated, the researchers explain in detail how the raters were trained and data collected, and advanced statistical techniques are employed. Surprisingly, it is one of very few published course evaluations in medical education employing the OSCE. Nevertheless, it provides an excellent example of what might be achieved in social work outcomes evaluation using this methodology.

## References

Badger, L. W. & MacNeil, G. (2002) 'Standardized clients in the classroom: a novel instructional technique for social work educators', *Research on Social Work Practice*, vol. 72, no. 3, pp. 364–374.

Baez, A. (2005) 'Development of an Objective Structured Clinical Examination (OSCE) for practicing substance abuse intervention competencies: an application in social work education', *Journal of Social Work Practice in the Addictions*, vol. 5, no. 3, pp. 3–20.

Bandura, A. (1989) 'Human agency in social cognitive theory', *American Psychologist*, vol. 44, pp. 1175–1184.

Bogo, M., Power, R., Regehr, C., Globerman, J. & Hughes, J. (2002) 'Evaluating a measure of student field performance in direct service: testing reliability and validity of explicit criteria', *Journal of Social Work Education*, vol. 38, no. 3, pp. 385–401.

Bogo, M., Regehr, C., Power, R., Hughes, J., Woodford, M. & Regehr, G. (2004) 'Toward new approaches for evaluating student field performance: tapping the implicit criteria used by experienced field instructors', *Journal of Social Work Education*, vol. 40, no. 3, pp. 417–426.

Carpenter, J. (2005) *Evaluating Outcomes in Social Work Education*, Social Care Institute for Excellence, London.

Cliff, N. (1987) *Analyzing Multivariate Data*, Harcourt Brace Jovanovich, San Diego, CA.

Council on Social Work Education (CSWE) (2008) *Educational Policy and Accreditation Standards*, CSWE, Alexandria, VA.

Cusimano, M. D., Rothman, A. & Keystone, J. (1998) 'Setting standards for performance assessment: defining standards of competent performance on an OSCE', *Academic Medicine*, vol. 73, no. 10, pp. 812–813.

Ellis, G. (2001) 'Looking at ourselves self-assessment and peer assessment: practice examples from New Zealand', *Reflective Practice*, vol. 2, no. 3, pp. 289–302.

Elstein, A., Shulman, L. S., Sprafka, S. A., Allal, L., Gordon, M., Jason, H., Kagan, N., Loupe, M. & Jordan, R. (1978) *Medical Problem Solving: An Analysis of Clinical Reasoning*, Harvard University Press, Cambridge, MA.

Gambrill, E. (2001) 'Educational policy and accreditation standards: do they work for clients?', *Journal of Social Work Education*, vol. 37, no. 2, pp. 226–240.

Harden, R. M. & Gleeson, F. A. (1979) 'Assessment of clinical competence using an Objective Structured Clinical Examination (OSCE)', *Medical Education*, vol. 13, pp. 41–54.

Hawker, J. A., Walker, K. Z., Barrington, V. & Andrianopoulos, N. (2010) 'Measuring the success of an objective structured clinical examination for dietetic students', *Journal of Human Nutrition and Dietetics*, vol. 23, pp. 212–216.

Herie, M. & Martin, G. W. (2002) 'Knowledge diffusion in social work: a new approach to bridging the gap', *Social Work*, vol. 47, no. 1, pp. 85–95.

Holden, G., Anastas, J., Meenaghan, T. & Metrey, G. (2002) 'Outcomes of social work education: the case for social work efficacy', *Journal of Social Work Education*, vol. 38, no. 1, pp. 115–133.

Koroloff, N. M. & Rhyne, C. (1989) 'Assessing student performance in field instruction', *Journal of Teaching in Social Work*, vol. 3, no. 2, pp. 3–16.

Lafrance, J., Gray, E. & Herbert, M. (2004) 'Gate-keeping for professional social work practice', *Social Work Education*, vol. 23, no. 3, pp. 325–340.

Leopold, Seth S., Morgan, Hannah D., Kadel, Nancy J., Gardner, Gregory C., Schaad, Douglas C. & Wolf, Fredric M. (2005) 'Impact of Educational Intervention on Confidence and Competence in the Performance of a Simple Surgical Task', *Journal of Bone and Joint Surgery*, vol. 87, no. 5, pp. 1031–1037.

Lozano, P., McPhillips, H., Hartzler, B., Robertson, A., Runkle, C., Scholz, K. A., Stout, J. W. & Kieckhefer, G. M. (2010) 'Randomized trial of teaching brief motivational interviewing to pediatric trainees to promote healthy behaviors in families', *Archives of Pediatrics & Adolescent Medicine*, vol. 164, no. 6, pp. 561–566.

Lu, Y. E., England, S., Feng, Y., Fong, R. & Yu, M. (2006) 'An alternative methodology in teaching and evaluating clinical competence', workshop presentation at the *52nd Annual Program Meeting*, Council of Social Work Education, Chicago, IL.

Maudsley, G. & Strivens, J. (2000) '"Science", "critical thinking" and "competence" for tomorrow's doctors. A review of terms and concepts', *Medical Education*, vol. 34, pp. 53–60.

Medina, C., Kwong, M. H. & Lu, Y. E. (2004) 'Assessing clinical competency in field education', in *Social Workers' Competency Assessment and Social Work Education Programs*, eds Y. Feng & Y. Lu, National University Press, Taipei, Taiwan.

Miller, M. (2004) 'Implementing standardized client education in a combined BSW and MSW program', *Journal of Social Work Education*, vol. 40, no. 1, pp. 87–102.

National Association of Social Workers (NASW) (2007) *Indicators for the Achievement of the NASW Standards for Cultural Competence in Social Work Practice*, NASW, Washington, DC.

Nisbet, J. & Shucksmith, J. (1984) *The Seventh Sense: Reflections on Learning to Learn* Scottish Council for Research in Education, Edinburgh.

Noble, J. H. & Stretch, J. J. (2002) 'Grade-induced beliefs about undergraduate generalist social work practice competency', *Evaluation Review*, vol. 26, no. 2, pp. 213–236.

O'Hare, T. & Collins, P. (1997) 'Development and validation of a scale for measuring social work practice skills', *Research on Social Work Practice*, vol. 7, no. 2, pp. 228–238.

O'Hare, T., Collins, P. & Walsh, T. (1998) 'Validation of the practice skills inventory with experienced clinical social workers', *Research on Social Work Practice*, vol. 8, no. 5, pp. 552–563.

Ogawa, T., Taguchi, N. & Sasahara, H. (2003) 'Assessing communication skills for medical interviews in a postgraduate clinical training course at Hiroshima University Dental Hospital', *European Journal of Dental Education*, vol. 7, no. 2, pp. 60–65.

Regehr, C., Regehr, G., Leeson, J. & Fusco, L. (2002) 'Setting priorities for learning in the field practicum. A comparative study of students and field instructors', *Journal of Social Work Education*, vol. 38, no. 1, pp. 55–66.

Regehr, G., Freeman, R., Hodges, B. & Russell, L. (1999) 'Assessing the generalizability of OSCE measure across continent domains', *Academic Medicine*, vol. 74, no. 12, pp. 1320–1322.

Reznick, R. K., Regehr, G., Yee, G., Rothman, A., Blackmore, D. & Dauphinee, D. (1998) 'High-stakes examinations: what do we know about measurement?', *Academic Medicine*, vol. 75, no. 10, pp. S97–S99.

Ryan, M., Habibis, D. & Craft, C. (1997) 'Guarding the gates of the profession: findings of a survey of gate-keeping mechanisms in Australian bachelor of social work programs', *Australian Social Work*, vol. 50, no. 3, pp. 5–12.

Scanlan, J. M. & Care, W. D. (2004) 'Grade inflation: should we be concerned?', *Journal of Nursing Education*, vol. 43, no. 10, pp. 475–478.

Shoemaker, J. K. & DeVos, M. (1999) 'Are we a gift shop? A perspective on grade inflation', *Journal of Nursing Education*, vol. 38, no. 9, pp. 394–398.

Strauss, A. (1987) *Qualitative Research for Social Scientists*, Cambridge University Press, Cambridge.

Tousignant, M. & DesMarchais, J. E. (2002) 'Accuracy of student self-assessment ability compared to their own performance in a problem-based learning medical program: a correlation study', *Advances in Health Sciences Education*, vol. 7, no. 1, pp. 19–27.

Turner, J. L. & Dankowski, M. (2008) 'Objective structured clinical exams: a critical review', *Family Medicine*, vol. 40, no. 8, pp. 574–578.

Ventimiglia, J. A., Marschke, J., Carmichael, P. & Loew, R. (2000) 'How do clinicians evaluate their practice effectiveness? A survey of clinical social workers', *Smith College Studies in Social Work*, vol. 70, no. 2, pp. 287–306.

Wilkinson, T. J., Frampton, C. M., Thompson-Fawcett, M. & Egan, T. (2003) 'Objectivity in objective structured clinical examinations: checklists are no substitute for examiner commitment', *Academic Medicine*, vol. 78, no. 2, pp. 219–223.

Ytterberg, S. R., Harris, I. B., Allen, S. S., Anderson, D. C., Kofron, P. M., Kvasnicka, J. H., McCord, J. P. & Moller, J. H. (1998) 'Clinical confidence and skills of medical students: use of an OSCE to enhance confidence in clinical skills', *Academic Medicine*, vol. 73, no. 10, pp. S103–S105.