An Evaluation of Multi-Station Objective Structured Clinical Examination (OSCE) in Clinical Psychology Training

Harriet Johnson (Clinical Psychologist), Kiki Mastroyannopoulou (Consultant Clinical Psychologist and Clinical Lecturer, University of East Anglia), Ella Beeson (Clinical Psychologist), Paul Fisher (Clinical Psychologist and Clinical Lecturer, University of East Anglia), and Margo Ononaiye (Clinical Psychologist and Deputy Clinical Director, University of Southampton)

harriet.johnson@nnuh.nhs.uk; Dr Harriet Johnson, Paediatric Clinical Psychology, West, Level 2, Norfolk and Norwich University Hospital, Colney Lane, Norwich, Norfolk, NR4 7UY.

This evaluation explored first year clinical psychology trainees' and assessors' experiences of Observed Structured Clinical Examination (OSCE). Changes to the OSCE based on their feedback helped reduce trainees' anxiety, promoted trainees' favourability of the OSCE and increased preparedness for placements.

Key words: clinical psychology, training, observed structured clinical examination

There is a need to improve assessment methods of clinical psychology trainees due to an overreliance on written assignments and supervisory assessment whilst on placement (Yap, Bearman, Thomas & Hay, 2012); the reliability and validity of these evaluations can be poor (Gonsalvez & Freestone, 2007; Scofield & Yoxtheimer, 1983). The British Psychological Society (BPS) (2014) requires clinical competencies of trainees to be assessed in vivo.

The Observed Structured Clinical Examination (OSCE) is routinely used to assess medical students' skills in interaction, diagnosis and intervention across standardised stations with trained actors as patients. It is recommended as a highly authentic, reliable and valid way of assessing competency (Kaslow et al., 2009). There has been little research into the use of competency-based assessments in clinical psychology (Roberts, Borden, Christiansen & Lopez, 2005).

Yap and colleagues (2012) explored trainees' views concerning the value and acceptability of the OSCE within clinical psychology training in Australia. They conducted a pilot study where nine participants completed the State-Trait Anxiety Inventory pre- and post-OSCE, a brief student feedback questionnaire and attended a focus group. Results showed that they viewed the OSCE as a valid, realistic, and fair assessment, despite high levels of anxiety pre-OSCE. The authors suggested that further evaluation of the validity of clinical psychology OSCE including assessors' opinions and larger sample sizes was needed.

The current study was an evaluation of the OSCE process introduced on the Clinical Psychology Doctoral Programme at the University of East Anglia for first year trainees. The OSCE assessed skills in clinical assessment, intervention, risk assessment and supervision on separate stations. Feedback from the 2014 cohort of trainees and assessors (Cohort One) informed changes to the 2015 OSCEs (Cohort Two). The two cohorts were given the same questionnaire in order to assess whether the changes improved the value and acceptability of the OSCE.

Method

Design

An independent measures mixed methods questionnaire design was used. Ethical approval was granted from the university. No personally identifiable information was collected. Data was stored according to university regulations.

Measure

Permission to use Yap et al.'s (2012) questionnaire was obtained. The questionnaire assessed trainees' perceptions of the "validity, relevance, realism, and fairness of the OSCE, as well as their experiences of anxiety and views on how well it fitted into clinical psychology training"

(Yap et al., 2012, p. 167). It comprised 10 items answered on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). It was adjusted so that it could also be completed by assessors; item 1 on the assessors' questionnaire related to the assessors' specific station. Both cohorts completed the same questionnaire.

Participants

In Cohort One, 13 out of 17 trainees and 18 out of 24 assessors completed the questionnaire. In Cohort Two, 20 out of 23 trainees and 30 out of the 35 assessors completed questionnaires. 33 trainees (82.5%) and 48 assessors (81%) participated in this study.

Procedure

Cohorts One and Two were given clear information in advance about the OSCE. The questionnaires were given to participants immediately after the OSCE and were collected by a member of staff. For Cohort One content analysis was used to identify the frequency of positive and negative comments to produce themes. These themes informed recommendations which were disseminated to the organisers of the OSCE. As a result the following changes were implemented for Cohort Two:

1. Trainee feedback forms were reworded to highlight "Areas of Development" and "Strengths" instead of "Appropriate/Learning Needs".
2. Course tutors emphasised that trainees would be expected to have several areas for development at this early stage of training.
3. Practical suggestions for development in relation to each learning need were added and three points for strengths and areas of development included in the summary section.

4. A Debrief session took place a week later with telephone support available in the interim.
5. Feedback was given by the university member of staff who assessed the trainee so that feedback could be contextualised. Feedback forms were uploaded to trainees' private student accounts.
6. The OSCE teaching session clarified that there were no expectations to prepare in advance.
7. The Supervision station was changed so that trainees could bring their own dilemma.
8. The Assessment station instructions were updated to clarify the nature/type of the assessment.
9. Briefing for assessors was increased in depth and clarified that trainees were encouraged to audio record the OSCE without needing to ask permission.
10. Course tutors clarified instructions regarding the need to bring notes.

The same process of content analysis was applied to feedback collected from Cohort Two. This process was conducted and reviewed by the authors collectively as with data collected from Cohort One. Independent samples t-tests, using the Statistical Package for the Social Sciences Version 22 (SPSS), were conducted to compare OSCE scores from both cohorts to evaluate the impact of the changes.

Results

Descriptive and inferential statistics

Visual inspection of the data suggested that there were differences between cohorts in how trainees and assessors rated the OSCE. All data from trainees and assessors were tested for

normality before running independent samples t-tests in SPSS. Infrequent missing data items were subject to mean imputation. All data were found to be normally distributed (i.e. the Shapiro-Wilk tests showed significance values greater than 0.5).

Tables 1 and 2 display the percentage of trainees and assessors who agreed or strongly agreed to the questionnaire items.

Table 1. Trainees' Questionnaire Results (Cohort One and Two)

Trainees	**Agree/Strongly Agree %**	
Questionnaire Item	**Cohort One**	**Cohort Two**
The OSCE appeared to provide a valid assessment of:		
a) Skills in engaging clients	77	80
b) Skills in assessing symptoms	77	65
c) Skills in assessing risk	92	80
d) Skills in diagnosis	8	5
e) Overall skills in clinical interviewing	77	90
2. The content of the stations was relevant to the course	85	100
3. The standardised clients were realistic	69	90
4. The OSCE was anxiety-provoking	100	75
5. The OSCE was more anxiety-provoking than any other types of examination	69	30
6. The OSCE was a fair assessment	31	75
7. I felt poorly prepared for the OSCE	15	25
8. The OSCE did not translate well to clinical practice	30	25

9. Overall the OSCE was a worthwhile exercise	77	90
10. Incorporating the OSCE into clinical training would be beneficial	69	85

Table 2. Assessors' Questionnaire Results (Cohort One and Two)

Assessors	**Agree/Strongly Agree %**	
Questionnaire Item	**Cohort One**	**Cohort Two**
1.The OSCE appeared to provide a valid assessment skills and knowledge relevant to my station	**88**	**100**
2.The content of the stations was relevant to the course	**89**	**93**
3. The standardised clients were realistic	**88**	**83**
4. The OSCE was anxiety-provoking	**94**	**56**
5. The OSCE was more anxiety-provoking than any other types of examination	**11**	**3**
6. The OSCE was a fair assessment	**95**	**96**
7. The trainees appeared poorly prepared for the OSCE	**0**	**0**
8. The OSCE did not translate well to clinical practice	**0**	**0**
9. Overall the OSCE was a worthwhile exercise	**100**	**97**
10. Incorporating the OSCE into clinical training would be beneficial	**94**	**97**

Overall scores

An independent samples t-test was conducted to compare trainees' OSCE scores between cohorts. There was a significant difference in the scores from Cohort One (M = 45.46, SD = 7.33) and Cohort Two (M = 51.35, SD = 5.71); t = (31) = -2.6, p = 0.015. This suggests that trainees in Cohort Two rated the OSCE more favourably.

An independent samples t-test was conducted to compare assessors' OSCE scores between cohorts. There was a non-significant difference in the scores for Cohort One (M = 40.50, SD = 2.60) and Cohort Two (M = 41.27, SD = 3.57); t = (46) = -.792, p = .432. This suggests that there was no significant difference between assessors' ratings of the OSCE.

Anxiety ratings

An independent samples t-test was conducted to compare trainees' ratings of anxiety between cohorts. Higher scores indicated greater anxiety. There was a significant difference in the scores from Cohort One (M = 4.69, SD = .480) and Cohort Two (M = 3.85, SD = .745); t = (31) = 3.61, p = .001. This suggests that trainees in Cohort Two found the OSCEs significantly less anxiety provoking.

An independent samples t-test was conducted to compare assessors' ratings of trainees' anxiety between cohorts. There was a non-significant difference in the scores from Cohort One (M = 4.11, SD = .583) and Cohort Two (M =3.67, SD = .994); t = (46) = 1.95, p = .057. This suggests that assessors in Cohort Two did not perceive a greater or lesser difference in trainees' anxiety.

Preparedness

Post-hoc analyses were conducted to determine whether trainees' feelings of preparedness differed between the two cohorts. Uncertainty concerning how much to prepare was a

prominent theme for Cohort One. Higher scores indicated that trainees felt less well prepared. There was a non-significant difference in the scores from Cohort One (M = 2.69, SD = .855) and Cohort Two (M = 2.60, SD = .883); t = (31) = .297, p = .768. This suggests that there was no difference in how well the trainees felt prepared for the OSCE.

However an independent samples t-test which was conducted to compare assessors' ratings of trainees' preparedness for the OSCE in Cohorts One and Two found a significant difference between the scores from Cohort One (M = 2.22, SD = .428) and Cohort Two (M = 1.67, SD = .606); t = (44.7) = 3.71, p = .001. This suggests that assessors perceived trainees in Cohort Two to be better prepared for the OSCE. The difference was found in the strength of disagreement with the statement (for the sake of brevity, disagreement scores are not shown in Tables 1 & 2).

Discussion

This study evaluated whether changes made to the OSCE for first year clinical psychology trainees made a difference to trainees' and assessors' perceptions of the value and acceptability of the OSCE.

Trainees in Cohort Two viewed the OSCE more favourably than trainees in Cohort One. The majority of trainees in Cohort Two agreed or strongly agreed that "incorporating the OSCE into clinical training would be beneficial" and "a worthwhile exercise" (85% and 90% respectively) compared with trainees in Cohort One (69% and 77%). A higher percentage of trainees in Cohort Two perceived the OSCE as "a fair assessment" compared with Cohort One (75% and 31% respectively). Many reasons could account for this. Firstly, the changes to the OSCE as a result of the recommendations may have improved the experience of the

OSCE for trainees. Secondly, Cohort Two reported significantly less anxiety which may have made the OSCE experience more positive. Thirdly, the fact that the OSCE had been undertaken by the previous cohort may have meant that trainees were primed to be more accepting of it.

No statistically significant change was found for assessors' perceptions of the OSCE. This may be because the recommendations were mainly focussed on improving the experiences of trainees. The majority of assessors in both cohorts viewed the OSCE as a valid and fair assessment and so may have reached a ceiling effect. The sensitivity and validity of the questionnaire to the recommendations is questionable. Yap et al.'s (2012) measure was designed to test whether trainees viewed the OSCE as a valid, fair and realistic assessment method, rather than measuring the specific recommendations generated in this evaluation.

Anxiety Ratings

Data analysis suggested that trainees in Cohort Two rated the OSCE as being significantly less anxiety provoking than trainees in Cohort One. A more consistent message to Cohort Two concerning preparation may have contributed to this. Due to the nature of the study's design it is not possible to tell which recommendations were most helpful in decreasing anxiety. There was no statistically significant change between the assessors' ratings of trainees' anxiety between the two Cohorts.

Preparedness

The results showed a significant difference between assessors' perceptions of trainees' preparedness across the two cohorts of trainees. Clearer guidance around preparation for the OSCE may have contributed to this. The non-significant difference between trainees' ratings

of preparedness might be because feeling prepared is very subjective and differs between individuals.

Conclusion

This report has evaluated the implementation of changes, informed by trainees' and assessors' feedback, to a university's Clinical Psychology Doctoral programme's OSCE process. The results showed that the second round of OSCEs (Cohort Two) were rated more favourably by trainees, were less anxiety provoking and trainees were better prepared according to assessors. The percentage of trainees and assessors who viewed the OSCE as a fair and valid assessment beneficial to clinical training increased following changes made to the OSCE. These results should be considered with caution since it remains unclear how much variance is explained by the recommendations as opposed to other factors, such as individual differences and cohort effects. Future evaluation of the university's OSCE could include a within-subjects component as the OSCE will be repeated for trainees in their second year. Focus groups would enable richer data to be collected to continue to improve the OSCE process.

References

Bagri, A. S., Zaw, K. M., Milanez, M. N., Palacios, J. C., Qadri, S. S., Bliss, L. A., & Ruiz, J. G. (2009). Geriatric medicine fellows' experiences and attitudes toward an objective structured clinical examination (OSCE). *Educational Gerontology*, *35*(4), 283-295.

Standards for Doctoral programmes in Clinical Psychology. (2014, May 1). Retrieved March 27, 2015, from http://www.bps.org.uk/system/files/Public files/PaCT/dclinpsy_standards_approved_may_2014.pdf/

Gonsalvez, C. J., & Freestone 1, J. (2007). Field supervisors' assessments of trainee performance: Are they reliable and valid?. *Australian Psychologist*, *42*(1), 23-32.

Kaslow, N. J., Grus, C. L., Campbell, L. F., Fouad, N. A., Hatcher, R. L., & Rodolfa, E. R. (2009). *Competency assessment toolkit for professional psychology. Training and Education in Professional Psychology, 3*(4 Suppl.) S27-S45.

Roberts, M. C., Borden, K. A., Christiansen, M. D., & Lopez, S. J. (2005). Fostering a Culture Shift: Assessment of Competence in the Education and Careers of Professional Psychologists. *Professional Psychology: Research and Practice*, *36*(4), 355.

Scofield, M. E., & Yoxtheimer, L. L. (1983). Psychometric issues in the assessment of clinical competencies. *Journal of Counselling Psychology*, *30*(3), 413.

Standards for Doctoral programmes in Clinical Psychology. (2014, May 1). Retrieved March 27, 2015, from http://www.bps.org.uk/system/files/Public files/PaCT/dclinpsy_standards_approved_may_2014.pdf/

Yap, K., Bearman, M., Thomas, N., & Hay, M (2012). Clinical Psychology Students' Experiences of a Pilot Objective Structured Clinical Examination. Australian Psychologist, 47(3), 165-173.

WORD COUNT = 2523