

REVIEW

Selecting doctors for postgraduate training in paediatrics using a competency based assessment centre

R Randall, H Davies, F Patterson, K Farrell

Arch Dis Child 2006;**91**:444–448. doi: 10.1136/adc.2005.076653

The design and implementation of an assessment centre in the South Yorkshire and South Humberside deanery for selecting doctors into postgraduate training in paediatric medicine is described. Eleven competency domains were identified in the job analysis. An assessment centre comprising of four exercises was implemented to assess candidates. There were modest relationships between candidates' performance on the various assessment centre exercises. Outcomes based on interview performance were related to, but not the same as, outcomes based on the combined results of the three other assessment centre exercises. Candidates perceived the assessment centre to be a fair selection method. It is concluded that an assessment centre approach to SHO recruitment is feasible and provides a greater breadth and depth of information about candidates than does a structured interview.

describe how it was used to develop an assessment centre as a replacement for interview-only assessment. We report on the logistics of running the assessment centre, and on the relationships between candidates' performance in a structured interview and their performance in other job related assessments. Data on candidates' perceptions of the assessment centre are presented.

METHODS

Identifying competency domains

A multi-source, multi-method job analysis was carried out using a similar methodology to that used to identify competency domains for general practitioners.⁶ This included four components: observation, by a trained occupational psychologist, of the practice of consultants (n = 4) over a total of 20 hours; three critical incidents focus groups with 25 doctors (8 consultants and 17 specialist registrars), and one with paediatric nurses (n = 4); and critical incidents interviews with patients (n = 31; mean age 8 years) and their parent(s) or carer(s) (n = 47). A total of 164 descriptions of behaviour were collected from these four components: 106 were indicative of good performance (positive indicators) and 58 of poor performance (negative indicators).

These indicators were each transferred onto a card for clustering into competency domains. Two independent pairs of occupational psychologists carried out a card sorting procedure which was the same as that used to identify competency domains for general practitioners.⁶ Inter-coder reliability was acceptable ($\kappa = 0.74$). Fourteen clusters (competency domains) were identified (table 1 and appendix 1 (www.archdischild.com/supplemental)).

An opportunity sample of 37 consultants in paediatrics (mean number of years in grade = 7.5 (SD 5.77)) was then asked to rate the importance of each competency domain with respect to its contribution to effective performance at SHO grade. The name of each competency domain and a summary of its behavioural indicators (appendix 1) were presented in a questionnaire. Participants rated the importance of each domain on a five point Likert-type scale (1 for little importance at SHO grade to 5 for a lot importance at SHO grade). Domains relating to interpersonal skills and integrity received the highest mean ratings (table 1). Two domains (teaching and managing others) were rated as having importance below the scale mid-point and were not targeted for assessment at selection.

We provided space on the questionnaire for participants to make comments about the

Research has shown that selection procedures can be improved through the use of a combination of competency based selection methods in an assessment centre.^{1–5} In the UK, competency based assessment centres have been used to select general practitioners and have shown good predictive validity.^{6,7} Modernising Medical Careers (MMC) emphasises the importance of robust processes for selection from Foundation to Speciality training programmes and this is an important area of work for all Royal Colleges.⁸ We report on the development and implementation of an assessment centre to select doctors for postgraduate training (senior house officer (SHO) posts) in paediatrics in the South Yorkshire and South Humberside (SYSH) deanery of the UK.

Assessment centres generally exhibit higher reliability and validity than selection interviews because they provide a more thorough and accurate assessment of the candidate.^{2,3} In addition, well executed assessment centres tend to be viewed as fair by candidates.^{9,10} These and other potential benefits, such as minimising failure at summative assessment,¹¹ can offset the costs of an assessment centre.⁵ However, for it to be effective, the content of an assessment centre must be developed from the results of a thorough job analysis.^{12,13} This identifies the competency domains that need to be assessed at selection.

We report on a job analysis for paediatric medicine carried out in the SYSH deanery, and



The appendices can be viewed on the ADC website (<http://www.archdischild.com/supplemental>)

See end of article for authors' affiliations

Correspondence to:
Dr H Davies, Sheffield Children's Trust, Western Bank, Sheffield S10 2TH, UK; h.davies@shf.ac.uk

Accepted 3 February 2006

Table 1 Competency domains identified in job analysis and examples of their behavioural indicators

Competency domain name	Mean of consultants' ratings importance at SHO grade (SD)
Professional integrity and respect for others (PIR)	4.40 (0.78)
Empathy and sensitivity (ES)	4.32 (0.80)
Personal attributes (such as flexibility and sense of humour) (PA)	4.18 (0.79)
Communication skills (CS)	4.13 (0.79)
Team work (TW)	4.10 (0.87)
Learning and personal development (LPD)	3.98 (0.80)
Coping with pressure (CP)	3.56 (0.97)
Personal organisation and administration skills (POA)	3.20 (0.99)
Vigilance and situational awareness (VSA)	3.17 (0.93)
Clinical/technical knowledge and expertise (CTK)	3.10 (0.81)
Conceptual thinking, problem solving, and decision making (CT)	2.95 (0.79)
Legal, ethical, and political awareness (LEP)	2.92 (1.05)
Managing others (MO)	2.38 (0.67)
Teaching (TE)	2.30 (0.76)

competency domains. Only three participants suggested modifications to the competency domains identified. As a result we modified some of the behavioural indicators (see table 1), but did not need to add other competencies. Almost all of the content of the personal attributes domain was moved into communication skills and empathy and sensitivity because of its overlap with these domains.

Development of selection methods

Three consultants in paediatrics from SYSH deanery developed seven tasks for piloting as assessment centre exercises. These were two group discussions, two written exercises, and three simulated consultations (two consultations with parents and one with a child and a parent). Current paediatric SHOs ($n=9$) and general practice trainees ($n=3$), all with less than six months' experience in post, participated in the piloting work. The exercise designers and three occupational psychologists observed and recorded participants' behaviour in the exercises. Two occupational psychologists then categorised the recorded examples of behaviour into the competency domains (see table 1 and appendix 1 (www.archdischild.com/supplemental)) to examine how well these exercises assessed the various domains. Feedback from participants was used to modify the instructions, and time allowed, for the exercises.

Assessment centre content and administration

Assessment exercises

The assessment centre consisted of the three exercises that, in the piloting work, had the most appropriate level of difficulty, and the strongest content validity and face validity. A structured interview, similar to that used in previous years' selection processes, was also used. Table 2 gives details of the assessment centre exercises and the competency domains they assessed. The more important competencies (table 1) were assessed more often in the assessment centre. A second simulated consultation (involving a child actor) has also been developed for use in future assessment centres.

Assessor training

All assessors received up to three hours face-to-face training (in small groups or individually) in the use of assessment materials. Training focused on the proper use of materials and included guidance on how to make reliable and valid

evaluations. Assessors were also provided with a training manual to support self-directed learning.

Assessment centre logistics

Ten trainee positions were available. Twenty seven candidates attended the assessment centre which lasted 8.5 hours. The four assessments were run in parallel and the order of exercises differed from candidate to candidate. A panel of three consultants carried out the interview. Each interviewer asked one question, with all three independently scoring every question.

Nine other trained assessors evaluated candidates. Five specialist registrars assessed the written exercise and the group exercise, and four consultants assessed the simulated consultation. All non-interview assessments were assessed on a one candidate to one assessor basis. Four or five candidates were assessed in each group exercise. Candidates completed the reflective written exercise immediately after completing the group exercise, with the same assessor evaluating both exercises. Each candidate was assessed by a total of five different assessors during the day.

Scoring of candidates and decision making

In the interview, each assessor scored each candidate out of 10 for each of three questions. A total interview score (maximum score 90) was calculated by summing the nine scores (three from each assessor). This was converted into a score from zero to nine (with one decimal point). In the three non-interview assessments, candidates were scored using the four stage process described in box 1 and appendix 2 (www.archdischild.com/supplemental).

An overall assessment centre (OAC) score was calculated. This was the sum of the overall exercise score for each exercise (a maximum of 21). This resulted in a 2.25 weighting for the interview score relative to each non-interview score, reflecting the fact that the score was made up of the assessments of three consultants. Candidates were ranked according to their OAC score. Then each candidate's performance was discussed in a session facilitated by the programme director and an occupational psychologist. Between-exercise consistencies and inconsistencies¹⁴ in competency domain scores were the focus of this discussion. Only those who had assessed the candidate contributed to the decision making process for a candidate. Final decisions were based on both OAC score and a qualitative discussion of the candidates' competency domain scores.

Candidate perceptions of the process

Candidates were asked to indicate their level of agreement with two statements ("The content of the selection centre seemed appropriate", and "I was given a good opportunity to show my skills and abilities today") on a five point Likert-type scale (from 1 for strongly disagree to 5 for strongly agree). Candidates were also asked to compare the assessment centre with other selection processes they had experienced during their medical career. The statement: "In comparison to other selection processes I have been through in my medical career..." was followed by two questions. The first question was: "The relevance of the content of the exercises to work in paediatrics was...". The second question was: "The level of opportunity I had to show my skills and abilities was...". Candidates responded to both questions on a seven point scale (1 for a lot less to 7 for a lot more).

Data analysis

All analysis was carried out using SPSS version 12. The deviation from normality of the distribution of exercise scores was examined using descriptive statistics. Correlations (Pearson's r) were used to assess whether candidates' performance in the interview was related to their perfor-

Table 2 Exercises used in the assessment centre and the competency domains assessed

Exercise	Summary of exercise content	Time	Competency domains												
			ES	CS	TW	CP	PIR	CT	VSA	POA	CTK	LPD	LEP		
Simulated consultation	<p>Consultation with the concerned parent (played by trained medical actor) of a 2 year old child</p> <p>Candidate is required to explore possible diagnoses and deal with the parent's questions/concerns</p> <p>Actor exhibits three emotions: fear, confusion, and frustration</p> <p>Actor delivers a series of prescribed key statements to express these three emotions</p>	20 minutes (5 minutes preparation and up to 15 minutes consultation)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Group exercise	<p>Discussion of and prioritisation of 11 competing tasks facing a paediatric SHO</p> <p>Tasks included administrative tasks, clinical tasks, discussions with patients, parents, and colleagues</p> <p>The group is required to decide on an order of priority for the tasks, and to provide a rationale for the prioritisation</p> <p>Assessment is based on content and conduct of the discussion as well as "correctness" of solution</p>	25 minutes (5 minutes individual preparation and up to 20 minutes group discussion)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Reflective written exercise	<p>Candidates present their own written response to the prioritisation task they dealt with in the group exercise (i.e. they rank the various tasks and provide their reasoning)</p> <p>Candidates give their personal reflection on their own performance and the performance of the group as a whole in the group exercise</p> <p>Assessment is based on the quality of their own response to the task, and their degree of insight into the group discussion</p>	25 minutes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Structured interview	<p>All candidates were asked the same three questions.</p> <p>Each question was designed to assess 2–3 competencies</p> <p>Candidates were asked: a question about their CV; one situational question about how they would deal with a risk management problem in paediatrics; and one question asking them to discuss how they had dealt with a particular situation (in relation to team working) in their medical career</p>	15 minutes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Box 1: Scoring process for non-interview assessments

During the exercise

- Stage 1: Observation of candidate
 - Assessors monitor both the verbal and non-verbal behaviour of one candidate
- Stage 2: Recording of candidate behaviour
 - Assessors make notes of what the candidate says and does as it happens. These are *factual observations*. Inferences are not made about what the behaviour means

After the exercise

(Stage 3 and Stage 4 take 15–20 minutes in total)

- Stage 3: Classification of observations
 - Assessors identify (within their recorded observations) behavioural indicators (both positive and negative) of the various competencies assessed by the exercise. Each observation is classified (e.g. a positive indicator of communication skills is allocated a code CS+). Assessors tally the positive and negative behaviours (on the scoring sheet) that they have observed for each competency
- Stage 4: Evaluation of candidate performance

Each competency is then given a score on a four point scale which is anchored as follows:

- 4 (good to excellent) = strong display of positive behavioural indicators
- 3 (satisfactory) = satisfactory display of positive behavioural indicators
- 2 (areas of concern) = many negative indicators displayed
- 1 (poor) = mostly negative indicators displayed

Considering the competency scores together, the assessors make a judgement on overall exercise performance. A score of 1 to 4 (on the scale above) is then given for the exercise overall, along with a brief justification of the overall score.

mance in other assessments. Two ranks were assigned to each candidate: one for the sum of non-interview assessment scores and one for the overall interview score. For each ranking a position outside of top 10 ranked candidates

(and ties) was coded as one. A position inside the top 10 ranked candidates (and ties) was coded as two. Then χ^2 analysis was used to test for a relationship between selection outcome based on interview alone and selection outcome based only non-interview assessments.¹⁵ The selection outcomes that were influenced by the inclusion of non-interview assessments were identified. Frequency analysis was used to evaluate the candidates' perceptions of the selection centre.

RESULTS

The distribution of all overall exercise scores was approximately normal (table 3). The mean scores for the interview and the group exercise were above the scale mid-point but not significantly skewed. These findings suggest that the exercises discriminated between candidates, but were neither too difficult nor too easy.

The structured interview score was significantly correlated with the group exercise score and the simulated consultation score (table 3). However, the size of these correlations indicated that the variability (r^2) in performance in the structured interview only accounted for 19.36% of the variability in simulated consultation scores and 22.10% of the variability in the group exercise. A total of 32.49% of the variability in reflective written exercise scores was accounted for by the variability in group exercise scores.

There was a significant relationship between the outcome based on structured interview scores only and the outcome based only on the sum of non-interview assessments ($\chi^2 (1,26) = 4.49$; $p < 0.05$). Nineteen of 27 candidates would have received the same outcome if it were determined by either of the two rankings alone. Eight out of the 11 highest scoring candidates at interview were among the top 13 scoring candidates based on the sum of non-interview scores (table 4). Three candidates who would have been selected based on interview performance alone were not offered a position because of their performance relative to other candidates on the non-interview assessments.

Candidates indicated that the assessment centre gave them either slightly more ($n = 5$), more ($n = 14$), or much more ($n = 6$) opportunity to demonstrate their abilities than other medical selection processes they had experienced. The vast majority indicated that the content of the selection centre was either more relevant to work in paediatrics (slightly more ($n = 4$), more ($n = 14$), and much more ($n = 6$)). Almost all either agreed ($n = 9$) or strongly agreed ($n = 15$) that the content of the selection centre seemed appropriate, and that they were given a good opportunity to show their skills and abilities at the assessment centre (14 agreed and 10 strongly agreed).

Table 3 Correlations between overall exercise scores

	Structured interview†	Group exercise	Written reflective exercise	Simulated consultation
Mean (SD)	60.10 (8.48)	3.01 (1.09)	2.52 (0.89)	2.67 (0.96)
Skew‡	-0.31	-0.65	0.16	-0.09
Kurtosis§	-0.30	-1.00	-0.60	-0.88
Inter-exercise correlations (r)				
Structured interview	-	0.47*	0.34	0.44*
Group exercise		-	0.57**	0.12
Written reflective exercise			-	0.21

n = 27.

*Significant at $p < 0.05$ (two tailed); **significant at $p < 0.01$ (two tailed).

†The mean inter-rater correlation of total interview scores was 0.89.

‡Standard error of skew = 0.45.

§Standard error of kurtosis = 0.87.

Table 4 Comparison of outcomes based on ranks for interview and ranks for sum of non-interview assessment methods

	Outcome based on sum of non-interview assessment		
	Not offer position	Offer position†	
Outcome based on structured interview (n)	Not offer position Offer position*	11 3	5 (3 in top 10 overall) 8‡(7 in top 10 overall)

*Includes two candidates tied for 10th and 11th rank.

†Includes five candidates tied for 9th, 10th, 11th, 12th, and 13th rank.

‡One candidate was among the ties for the lowest rank in ties in both interview and non-interview assessment.

DISCUSSION

There is considerable pressure to develop more robust methods for selection as part of the implementation of MMC. Candidates applying for postgraduate training in paediatric medicine should be assessed across a wide range of competency domains and in a variety of situations. Interpersonal skills, team working skills, professional integrity, an insight into one's own skills and development needs, and an ability to cope with pressure were all identified as important skills by the job analysis and should all be assessed at selection. An assessment centre provides a way of assessing candidates' aptitude in these domains in several contexts such as a team environment, in an individual consultation, or when dealing with written tasks.

Many candidates performed consistently across interview and non-interview based assessments. However, some who did well at interview (where scores were based on candidates' self-reports) did less well in job related tasks (where scores were based on their demonstration of their skills) and vice versa. The correlation between performance in the group exercise and the simulated consultation was also low. An assessment centre allows for consistency in performance across situations to be considered during decision making.

Candidate perceptions of the fairness of the assessment centre were extremely positive, and they saw it as fairer than other selection processes they had experienced. Previous research indicates that perceived fairness is related to favourable perceptions of the recruiting organisation.^{9, 10} Many candidates commented that the assessment centre had helped them to learn something about their strengths and weaknesses. Assessment centre results can be used for identifying the immediate training needs of new employees.¹⁶

The assessment of 27 candidates in one day shows that high volume recruitment is possible using an assessment centre. As assessors become skilled in more than one exercise, the assessor:candidate ratio will become approximately 1:3.

We present data which were drawn from only one area of the UK and our findings need to be validated on a wider sample. Additional exercises, including a simulated consultation, which includes a child and an assessment of technical aptitude could be included. As well as further evaluation of this approach on a wider sample, determination of predictive validity in terms of future competence is essential. The

competencies we have identified should be used to assess trainees' progress (both during training and at summative assessment) and to structure feedback for them. Further work exploring how best to utilise the output from the selection process to inform personal development planning is also important. However, our results from this initial pilot suggest that this approach to selection is worth pursuing and suggest direction for further work.

ACKNOWLEDGEMENTS

The ethics committee of City University Psychology department approved this work. Dr Kath Farrell collected and analysed the data for the job analysis phase of the research. Dr Gail Moss, Dr Carrie Mackenzie, Dr Julian Archer, and Mr Jonathan Beard contributed to the development of the assessment centre exercises, and Ms Corina Voelklein helped to produce the assessment materials and administer the assessment centre. We would also like to the doctors who worked as assessors.

Authors' affiliations

R Randall, F Patterson, K Farrell, Department of Psychology, City University, London, UK

H Davies, Sheffield Children's Trust, Sheffield, UK

Funding: The postgraduate dean (Dr Sarah Thomas) of the South Yorkshire and South Humber Deanery funded the development and implementation of the assessment centre, and has supported the decision to seek publication in this journal. We are grateful to her for her support.

Competing interests: none declared

REFERENCES

- 1 Cook M. *Personnel selection: adding value through people*, 3rd edn. Chichester: John Wiley and Sons, 1998:173–92.
- 2 Hunter J, Hunter R. Validity and utility of alternative predictors of job performance. *Psychol Bull* 1984;**96**:72–98.
- 3 Robertson IT, Smith M. Personnel selection. *J Occup Organ Psych* 2001;**74**:441–72.
- 4 Salgado J, Viswesvaran C, Ones D. Predictors used for personnel selection: an overview of constructs, methods and techniques. In: Anderson N, Ones D, Sinangil H, et al. *Handbook of Industrial, Work and Organizational Psychology, Volume 1*. Thousand Oaks, CA: Sage, 2001:165–99.
- 5 Schmidt FL, Hunter JE. The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psyc Bull* 1998;**124**:262–74.
- 6 Patterson F, Ferguson E, Lane PW, et al. Competency model for general practice: implications for selection, training and development. *Br J Gen Pract* 2000;**50**:188–93.
- 7 Patterson F, Ferguson E, Norfolk T, et al. A new selection system to recruit general practice registrars: preliminary findings from a validation study. *BMJ* 2005;**330**:711–14.
- 8 Department of Health, Social Services and Public Safety. *Modernising medical careers: the response of the four UK Health Ministers to the consultation on "Unfinished business—proposals for reform of the senior house officer grade"*. London: Department of Health, 2003.
- 9 Macan TH, Avedon MJ, Paese M, et al. The effects of applicants' reactions to cognitive ability tests and an assessment centre. *Pers Psych* 1994;**43**:615–31.
- 10 Rynes SL, Connerley MR. Applicant reactions to alternative selection procedures. *J Bus Psychol* 1993;**7**:261–77.
- 11 Crawford ME. Commentary: Reassuring evidence on competency based selection. *BMJ* 2005;**330**:714.
- 12 Algera J, Grueter MAM. Job analysis for personnel selection. In: Smith M, Robertson I, eds. *Advances in selection and assessment*. Chichester: John Wiley and Sons, 1989:7–30.
- 13 Pearn M, Kandola R. *Job analysis: a manager's guide*. London: Institute of Personnel Management, 1993.
- 14 Bycio P, Hahn J, Alvares K. Situational specificity in assessment center ratings: a confirmatory factor analysis. *J Appl Psych* 1987;**72**:463–74.
- 15 Tabachnick BG, Fidell LS. *Multivariate statistics*, 4th edn. Boston: Allyn and Bacon, 2001:55.
- 16 Hough L, Oswald FL. Personnel selection: looking toward the future—remembering the past. *Ann Rev Psychol* 2000;**51**:631–64.