

Clinical scenario-based assessments to measure reasoning in complex cases: Are we fully realising their potential in specialty training?

*by Zaid Ali, Reza Vahid Roudsari and Julian Satterthwaite*

Planning the management of complex cases requires years of structured training, with experienced gained both by attending joint/multidisciplinary clinics and through managing patients needing interdisciplinary care. Together, these reinforce basic skills and develop high-order thinking. It is crucial, however, that the development of these advanced planning skills can be measured, so as to assess the progress of trainees and the effectiveness of training programmes. Here, the validity, reliability and consistency of a clinical, hypodontia case-based structured assessment tool of clinical reasoning and higher-order thinking is tested.

*Authors:* Mr Zaid Ali\*, Specialty Registrar (Academic) in Restorative Dentistry, Charles Clifford Dental Hospital, Sheffield, Mr Reza Vahid Roudsari, Clinical Lecturer and Honorary Specialist Registrar in Restorative Dentistry, School of Dentistry, University of Manchester and Professor Julian D Satterthwaite, Professor and Honorary Consultant in Restorative Dentistry, School of Dentistry, University of Manchester.

\*Corresponding author  
E: Zaid.ali@sheffield.ac.uk

*Keywords:* Clinical reasoning, hypodontia, dental education, measurement of clinical reasoning, case-based assessment, psychometric tools



## Introduction

Joint/multidisciplinary clinics for the planning of particularly complex cases are well established in many UK teaching and district general hospitals. Cases of cleft lip and palate, trauma and conditions affecting the number, condition or position of teeth, as well as cancer cases, are jointly planned by clinicians with various specialty interests and skillsets to reach a mutually agreed, realistic endpoint. To acquire the skills needed as a specialist in any one of the dental specialties requires years of structured training. Trainees gain experience in treatment-planning complex cases, perhaps by attending joint specialty or multidisciplinary team (MDT) treatment planning clinics, and managing patients with complex interdisciplinary care requirements.<sup>1</sup> This is thought to reinforce basic skills and develop higher-order thinking in deciding which of several options are most appropriate for patients with complex conditions and challenging presentations.

Hypodontia, or the congenital absence of permanent teeth, has a prevalence of 1–11.3% and is one such condition.<sup>2,3</sup> Patients present with varying degrees of severity and functional and aesthetic problems, and, as such, pose challenges for treatment planning and management to achieve good, stable, long-term function and, in many cases, aesthetics.<sup>3</sup> Treatment may involve orthodontics to close or open spaces, direct or indirect restorative techniques to change tooth shape or occlusal vertical dimensions, and fixed or removable prostheses to restore space.<sup>3</sup> Variations in skeletal growth patterns may also cause an exaggerated mismatch between the maxillary and mandibular arches, potentially corrected by orthognathic surgery. Although multidisciplinary care is not always available, input from various specialties is recommended at an early stage to ensure all options are considered carefully.<sup>4,5</sup>

Complex cases challenge the clinical reasoning and treatment planning skills of even expert clinicians. Measuring clinical reasoning in a wider sense is of great importance in the context of measuring trainee progress and certification of training. A number of tools are used in dental and medical education for measuring knowledge, clinical reasoning and mental dexterity. Those relevant to clinical reasoning include objective structured clinical examinations (OSCE),<sup>6,7</sup> clinical reasoning problems (CRPs);<sup>8</sup> key features problems (KFPs)<sup>9</sup> and; script concordance tests (SCTs).<sup>10</sup>

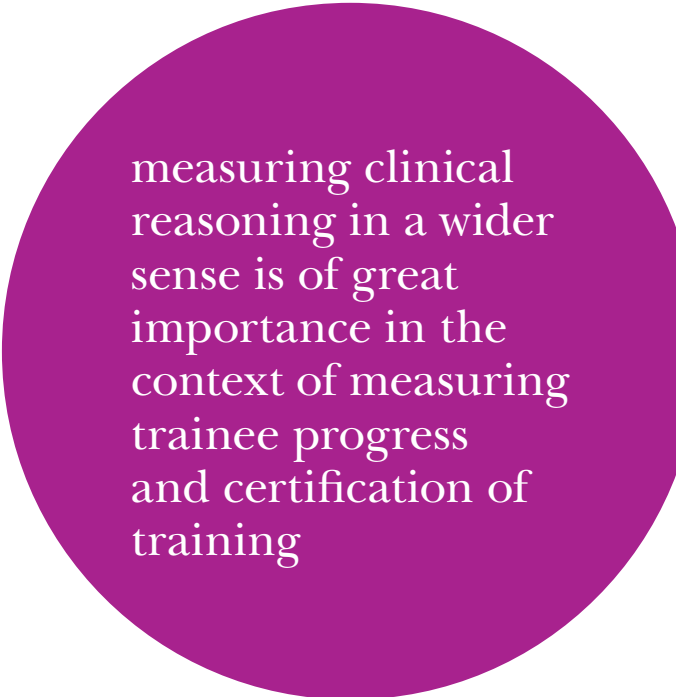
Measuring and interpreting the validity of these tools has been reviewed extensively in the literature.<sup>11,12</sup> In restorative dentistry, a case-based structured oral examination is used to measure candidate's clinical reasoning and higher-order thinking. During training, a number

of other tools, known as workplace-based assessments (WBAs), are used to measure progress. Clinical case-based scenarios, as a mock-exam style assessment, may be another, informal way to measure trainee progress and prepare them for the intercollegiate specialty fellowship examination (ISFE), which is the exit examination for specialty trainees in the surgical specialties.

We investigated the construct validity, inter-rater reliability and internal consistency of a clinical, hypodontia case-based structured assessment tool of clinical reasoning and higher order treatment planning in UK restorative dentistry trainees.

## Materials and methods

Trainees were identified by contacting members of the Specialist Registrars in Restorative Dentistry Group (SRRDG) and delegates attending the SRRDG 2012 conference. Trainees were included if they were in a restorative dentistry specialty registrar/specialist registrar post and had a national training number. No incentives were used and participation was voluntary. Registrar year of training and experience of hyperdontia MDT clinics was determined from a pre-assessment questionnaire.

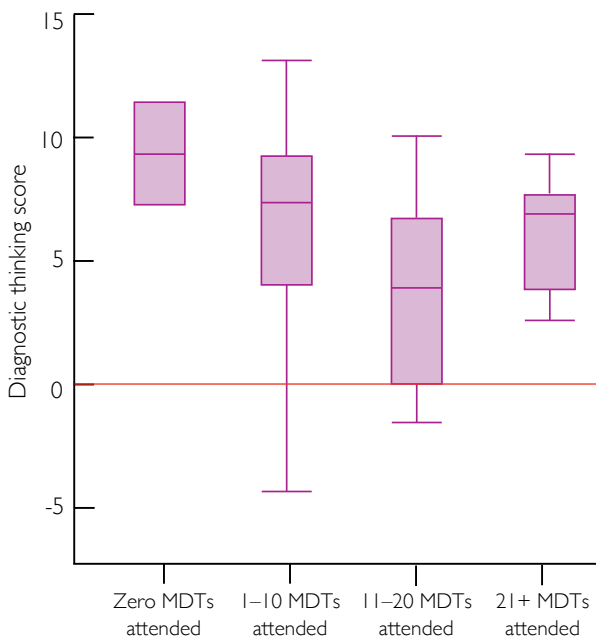


measuring clinical reasoning in a wider sense is of great importance in the context of measuring trainee progress and certification of training

**Table 1** Number of trainees by year of training and number of MDT clinics attended.

| N=26                                  |    |
|---------------------------------------|----|
| <b>Year of training</b>               |    |
| 1                                     | 3  |
| 2                                     | 3  |
| 3                                     | 5  |
| 4                                     | 7  |
| 5                                     | 6  |
| Not answered                          | 2  |
| <b>Number of MDT clinics attended</b> |    |
| Zero                                  | 2  |
| 1–10                                  | 11 |
| 11–20                                 | 6  |
| 21+                                   | 7  |

**Figure 1** Box and whisker plot showing median, interquartile range and range of diagnostic thinking scores, by number of MDTs attended.



A written assessment tool was developed to test the clinical reasoning skills of restorative dentistry trainees by assessing their responses to clinically oriented questions using real hypodontia case material. Three cases were selected: one example case and two assessment cases. A detailed case history, with clinical findings, special investigations, clinical photographs, radiographs and orthodontic advice, was presented. All cases were from the Joint Hypodontia Clinic of the University Dental Hospital of Manchester. Consent for the use of clinical photographs and non-identifiable material was obtained for all cases.

Features of several clinical reasoning measurement tools were incorporated into the assessment: 'key features' from KFPs<sup>9</sup> and CRPs (CRPs);<sup>7</sup> 'free text' similar to OSCE post-encounter forms and the justification of management decisions in CRPs;<sup>7,8</sup> and 'decision making' from SCTs.<sup>11</sup> They measured the ability to identify factors affecting the treatment plan; barriers to success; and further information to form and outline a treatment plan, explaining its strengths and limitations.

The assessment was piloted on four restorative dentistry registrars. A clinical lecturer (RVR) and professor in restorative dentistry (JDS) reviewed the assessments. Model answers and pass marks, at the level of ISFE exit examination, were developed from the responses of five consultants in restorative dentistry, following the approach outlined by Elstein *et al.*<sup>13,14</sup> and used in the SCT.<sup>10,13,14</sup> Two assessors (ZA and RVR) marked the scripts. The maximum score for each question was determined following the inspections of expert responses, with an overall score of 70. The marking was calibrated using a sample of five randomly selected scripts, and areas of contention were discussed and reconciled prior to individual marking. The assessors then developed a 'consensus' mark. Where consensus could not be reached, the third author (JDS) gave the final verdict. Outcome scores for each item were calculated relative to the pass mark, with 0 a pass, positive scores above a pass and negative scores below a pass. Item scores were combined to give a total score relative to the pass mark.

Trainees were also asked about how the training programmes could be improved in terms of planning complex cases, and any themes identified.

The University of Manchester Ethics Committee and the Research Conduct and Accountability Committee exempted the study from ethical review.

#### Statistical analysis

Inter-examiner reliability was investigated using the intraclass correlation coefficient (ICC). Internal consistency was measured using Cronbach's alpha. Outcomes were presented with descriptive data, including the median and interquartile range and range for each trainee category. The Kruskal-Wallis test was used to compare the trainees' 'diagnostic thinking' and 'total clinical reasoning/higher-order thinking' scores, grouped by the number of MDT meetings attended and year of

training. SPSS for Windows (version 16.0, SPSS Inc. Chicago, IL, USA) was used to perform the analyses.  $p < 0.05$  was considered statistically significant.

## Results

Sixty-three trainees were invited to participate, with 26 scripts completed, at a response rate of 41.3%. Seven respondents (26.9%) were academic, and 19 (73.1%) NHS trainees. Year of training and experience of MDT clinics is shown in Table 1.

Inter-rater reliability of the assessment was excellent, with an ICC of 0.901. Combining the all items into one clinical reasoning or 'higher-order thinking' score gave Cronbach's alpha of 0.889.

Figures 1–4 show the box and whisker plots of the 'diagnostic thinking' and 'total clinical reasoning/higher-order thinking', by MDT experience and year of training. The number of MDT meetings attended and the year of training did not significantly affect 'diagnostic thinking' ( $p = 0.28$  and  $p = 0.09$  respectively) or 'total clinical reasoning/higher-order thinking' ( $p = 0.95$  and  $p = 0.98$  respectively) scores. Nevertheless, high variability in scores was observed for respondents in the early years of training and with less experience of MDTs. The variability around the pass mark reduced with increasing years of training and experience of MDTs, and converged around the pass mark.

The majority of trainees felt that their programme fulfilled their requirements in terms of managing complex cases, with only 11.5% ( $n = 3$ ) disagreeing. From the responses to the question 'What changes would you like to see in your program to improve the training in treatment planning for hypodontia cases?', three main themes emerged: increased didactic teaching; increased interaction with other specialties; and protected clinical training.

## Discussion

The assessment tool was developed to measure clinical reasoning or higher-order thinking in restorative dentistry trainees. These are complex constructs to measure and we neither suggest that clinical case-based assessment tools have fully explored and measured them in each case, nor that written scenario-based examination should replace the measures used in high-stakes examinations such as the ISFE. We made every effort to make the tool clinically representative, using good-quality clinical photographs and radiographs of a real patient, and providing additional information, including the opinions of an orthodontist.

Although it is not possible to entirely prove the construct validity of the assessment tool in this way, concepts from previously validated tools measuring similar properties were incorporated. The identification of important features and justifying a treatment plan is similar items to that in CRPs.<sup>8</sup> Trainees were also measured against recognised experts in the field (ie consultants in restorative dentistry), which is a tried-and-tested technique.<sup>10,13,14</sup> Features from CREs, the

Figure 2 Box and whisker plot showing median, interquartile range and range of total higher-order thinking scores, by number of MDTs attended.

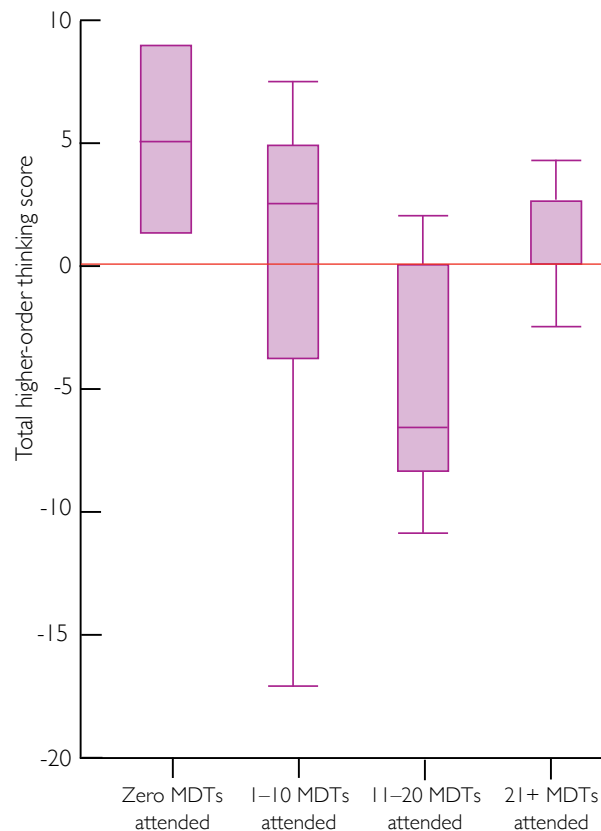


Figure 3 Box and whisker plot showing median, interquartile range and range of diagnostic thinking scores by year of specialty training.

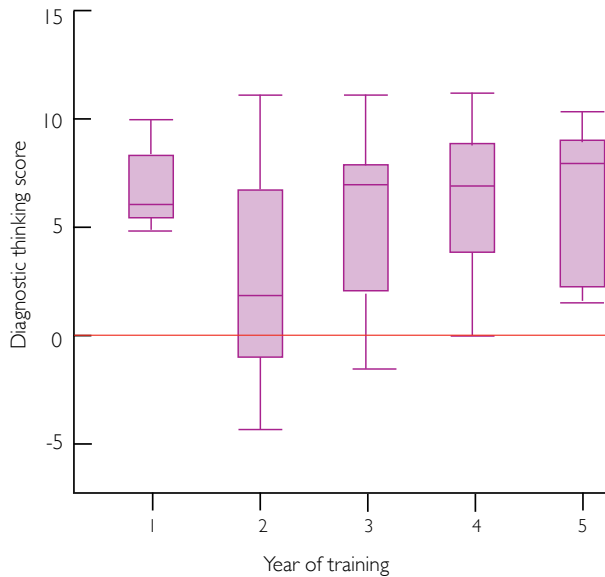
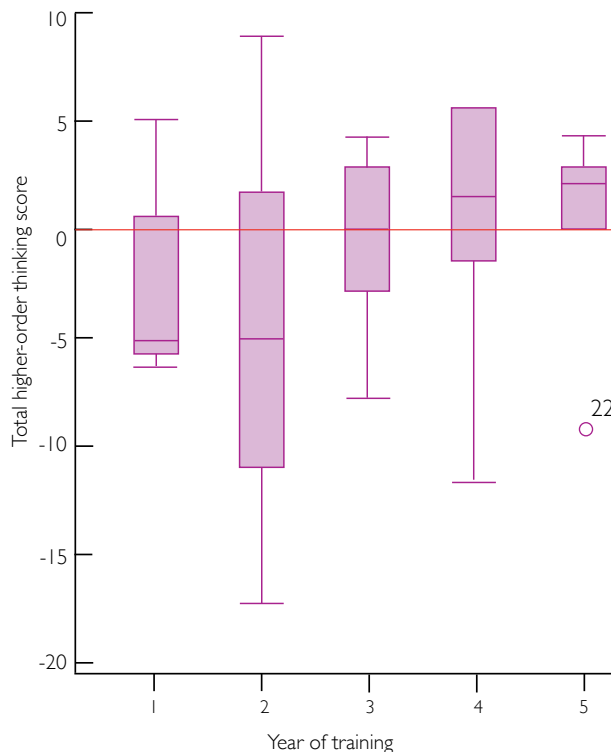


Figure 4 Box and whisker plot showing median, interquartile range and range of total higher-order thinking scores by year of specialty training.



assessment of treatment planning decisions from OSCE and oral (viva voce) examinations and decision-making skills similar to those tested in the SCT were also incorporated to improve validity.<sup>8,10,15</sup> We accept that we have assumed combining these various items would give the assessment tool a degree of validity in measuring diagnostic ability and clinical reasoning. Although this requires further validation, the assessment showed good internal consistency, thus justifying our approach. The inter-rater reliability was excellent, supporting the scoring method and the reliability of the results.

With a small sample of only 26 trainees' data could only be presented descriptively. With no power calculation or estimate of effect size, the study was effectively an unpowered observational study. Specialty registrars generally are committed to a number of roles alongside their clinical training, and taking significant time out to consider cases carefully and give full answers was not a priority. Incorporating mock clinical exams as part of regional and national study days may be one means of measuring the outcome of larger numbers of trainees and, hence, improving response rates. Furthermore, only eight items from two clinical cases were used, which is small compared with tools used for medical trainees, such as Script Concordance Tests, which recommend a minimum of 20 cases to ensure the validity of the findings. This would not have been feasible in this study, as we relied on voluntary compliance, and increasing the number of questions would have risked further reducing the response rate.

Other sources of bias may include the lack of standardisation in time allowed to complete the assessment, no account taken of trainees prior experiences and qualifications, and no objective measure for the quality of training experienced at the trainees' respective dental schools or hospitals.

The ability to measure reliably and accurately a clinician's clinical reasoning remains a challenge. In an age where the revalidation of medical and dental professionals is the focus of increased attention, the development of tools and measures may be increasingly useful. Informal assessment using clinical scenario-based tools may help measure trainees' progress through a specialty training programme alongside other measures, such as WBAs and the Annual Record of Competency Progression.

### Conclusion

We have demonstrated good inter-rater reliability and internal consistency with a clinical case based assessment tool to measure clinical reasoning. Although we cannot say with any certainty how many MDTs a trainee should attend, there was reduced variability in candidates who had attended 21 or more MDTs and were in their final year of training. Further research is required to determine if such tests are able to demonstrate good intra-rater, test-retest reliability, and compare with other accepted measures, such as OSCEs or oral assessments.

### Acknowledgments

The authors would like to thank the trainees who volunteered to take part and the consultant expert clinicians who kindly helped in the development of model answers: Martin Ashley, David Eldridge, Ian Harris, Rajesh Patel and Professor Julian Satterthwaite.

### References

1. Specialist Advisory Committee in Restorative Dentistry. *Curriculum for Specialty Training In Restorative Dentistry*. Approved by the GDC April 2009: General Dental Council; 2009.
2. Larmour C, Mossey P, Thind B *et al*. Hypodontia--a retrospective review of prevalence and etiology. *Part I. Quintessence Int* 2005; **36**: 263–270.
3. Bishop K, Addy L, Knox J. Modern restorative management of patients with congenitally missing teeth: 1. Introduction, terminology and epidemiology. *Dent Update* 2006; **33**: 531–534, 537.
4. Addy L, Bishop K, Knox J. Modern restorative management of patients with congenitally missing teeth: 2. Orthodontic and restorative considerations. *Dent Update* 2006; **33**: 592–595.
5. Worsaae N, Jensen BN, Holm B, Holsko J. Treatment of severe hypodontia-oligodontia--an interdisciplinary concept. *Int J Oral Maxillofac Surg* 2007; **36**: 473–480.
6. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Medical Education* 1979; **13**: 41–54.
7. Durning SJ, Artino A, Boulet J *et al*. The feasibility, reliability, and validity of a post-encounter form for evaluating clinical reasoning. *Medical teacher* 2012; **34**: 30–37.
8. Groves M, Scott I, Alexander H. Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. *Medical Teacher* 2002; **24**: 507–515.
9. Doucet MD, Purdy RA, Kaufman DM, Langille DB. Comparison of problem-based learning and lecture format in continuing medical education on headache diagnosis and management. *Medical Education* 1998; **32**: 590–596.
10. Charlin B, Roy L, Brailovsky C *et al*. The Script Concordance test: a tool to assess the reflective clinician. *Teaching and Learning in Medicine* 2000; **12**: 189–195.
11. van der Vleuten CP, Newble DI. How can we test clinical reasoning? *Lancet* 1995; **345**: 1,032–1,034.
12. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *The American Journal of Medicine* 2006; **119**: 166.e167–116.
13. Elstein AS, Shulman LS, Sprafka SA. Medical problem-solving. *Journal of Medical Education* 1981; **56**: 75–76.
14. Elstein AS, Shulman LS, Sprafka SA. Medical Problem Solving: A Ten-Year Retrospective. *Evaluation and the Health Professions* 1990; **13**: 5–36.
15. Wood TJ, Cunnington JP, Norman GR. Assessing the measurement properties of a clinical reasoning exercise. *Teaching and Learning in Medicine* 2000; **12**: 196–200.