

Early assessment with a virtual reality haptic simulator predicts performance in clinical practice

Loulwa M Al-Saud ^{1,2}, Faisal Mushtaq ³, Richard P Mann,⁴ Isra'a Mirghani,⁵ Ahmed Balkhoyor,^{5,6} Richard Harris,³ Cecilie Osnes ⁷, Andrew Keeling,⁷ Mark A Mon-Williams,³ Michael Manogue⁷

¹Operative Division, Department of Restorative Dental Sciences, King Saud University College of Dentistry, Riyadh, Saudi Arabia

²School of Dentistry, University of Leeds, Leeds, UK

³School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds, UK

⁴School of Mathematics, Faculty of Mathematics and Physical Sciences, University of Leeds, Leeds, UK

⁵School of Dentistry and School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds, UK

⁶Department of Preventive Dentistry, Faculty of Dentistry, Umm Al-Qura University, Makkah, Saudi Arabia

⁷School of Dentistry, Faculty of Medicine and Health, University of Leeds, Leeds, UK

Correspondence to

Dr Faisal Mushtaq, School of Psychology, Faculty of Medicine & Health, University of Leeds, Leeds LS2 9JT, UK; pscicon@leeds.ac.uk

Accepted 28 September 2019

ABSTRACT

Background Prediction of clinical training aptitude in medicine and dentistry is largely driven by measures of a student's intellectual capabilities. The measurement of sensorimotor ability has lagged behind, despite being a key constraint for safe and efficient practice in procedure-based medical specialties. Virtual reality (VR) haptic simulators, systems able to provide objective measures of sensorimotor performance, are beginning to establish their utility in facilitating sensorimotor skill acquisition, and it is possible that they may also inform the prediction of clinical performance.

Methods A retrospective cohort study examined the relationship between student performance on a haptic VR simulator in the second year of undergraduate dental study with subsequent clinic performance involving patients 2 years later. The predictive ability was tested against a phantom-head crown test (a traditional preclinical dental assessment, in the third year of study).

Results VR scores averaged across the year explained 14% of variance in clinic performance, while the traditional test explained 5%. Students who scored highly on this averaged measure were ~10 times more likely to be high performers in the clinical crown test. Exploratory analysis indicated that single-trial VR scores did not correlate with real-world performance, but the relationship was statistically significant and strongest in the first half of the year and weakened over time.

Conclusions The data demonstrate the potential of a VR haptic simulator to predict clinical performance and open up the possibility of taking a data-driven approach to identifying individuals who could benefit from support in the early stages of training.

INTRODUCTION

Prediction of future clinical performance is a fundamental concern for medical educators. Early identification of likely successful students has pedagogical, administrative, patient safety and economic implications.^{1–2} Tests that are able to predict performance would allow for a more robust selection process and could inform early interventions for students struggling with the demands of the programme.^{3–5}

The successful practice of procedure-based medical and dental specialties such as surgery and restorative dentistry requires an individual to display high levels of academic and sensorimotor ability in tandem. Making predictions about the former has been a relatively easy task (though still fraught with difficulties) in comparison with the latter. Historically, the approach to measuring a student's intellectual aptitude

has involved the use of data from standardised national academic and psychometric tests (examples include, but are not limited to, grade point averages,⁶ the dental aptitude test,^{7–8} the United Kingdom Clinical Aptitude Test⁹ and personality profile approaches.^{10–11} The base-level admissions criteria for dentistry within a country are generally consistent across schools, with little variation. In contrast, measuring the functional, perceptual and sensorimotor capabilities of a student has proven to be much more challenging, and there is no consensus on what types of capabilities need to be measured and how they might relate to dentistry.¹² Measurement of these abilities varies substantially across dental schools.

To date, a wide range of predictors have been investigated. Tests include task-based approaches such as chalk carving,^{13–14} waxing,¹⁵ wire bending,^{16–17} using tweezers¹⁸ and manipulating small parts.¹⁹ Others involve more general measures of traits that could potentially capture dentistry-relevant skills, including spatial ability,²⁰ broad psychometric tests²¹ and other predictors.^{22–23} The majority of these approaches have had limited predictive value, with a call to use these as screening tools instead.¹ Unfortunately, such approaches are ethically problematic if there is no empirical evidence to show a relationship between the measure and the ability to practice dentistry.

One potential reason for the low level of predictive value is that we have a very limited understanding of the fundamental traits that the tests are capturing and how these relate to specific dental tasks.²¹ This has led some researchers to use dentistry-specific tasks in preclinical settings to predict clinical performance, reasoning that the preclinical simulated restorative tasks share a large degree of overlap with restorative procedures used in the clinic. The success of these approaches varies considerably, with evidence ranging from moderate success²⁴ through weak predictive value²⁵ to no relationship.²⁶

Recent advances in technology have led to the increasing prevalence of virtual reality (VR) simulators in dental education. A body of accumulated evidence supports the validity²⁷ of many VR simulators, such as discriminant evidence (ie, ability of simulators to discriminate between novice and accomplished users)²⁸ and content-based validity evidence.^{29–32} The systems are increasingly being adopted to complement traditional training approaches, and there is growing evidence of their value in facilitating sensorimotor skill acquisition.³³ The fact that these systems are able to provide



© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Al-Saud LM, Mushtaq F, Mann RP, *et al.* *BMJ Stel* 2020;**6**:274–278.

precise measures of sensorimotor performance relevant to real-world dental performance suggests that performance on these systems could be beneficial in predicting subsequent clinical aptitude. Some early studies involving these VR systems have shown that pretest simulator performance correlates with early preclinical course performance but not later performance where more complex dental procedures were involved.³⁴ Another study demonstrated performance on a VR simulator correlated positively and predicted performance in a preclinical manikin course.³⁵ Further investigations have shown that pretest performance is a strong predictor of early but not late preclinical operative dentistry performance.³⁶

A new generation of VR systems incorporating haptic technology has provided a step change in dental simulation. Their ability to precisely capture kinematic information, while participants complete dentistry-specific tasks, offers an opportunity to deliver objective performance metrics, and these systems have shown substantial promise in prediction. For example, performance on a basic haptic exercise was found to be a reasonable predictor of students' performance in a preclinical operative dentistry course.³⁷ In related previous research, three haptic dental tasks were compared to identify the best predictors of preclinical operative dentistry performance. Strong associations were found between performance on complex haptic exercises and preclinical operative dentistry performance.^{2 12}

There is growing evidence that haptic simulators may be able to provide measures of student performance that can predict subsequent dental performance, but this work has thus far been focused only on preclinical performance. In this study, we attempted to address this gap by examining how well a variety of manual dexterity tasks on a haptic simulator could predict students' performance on a crown preparation in the clinic. We contrast the VR predictive ability with crown preparation on a 'typodont' in a traditional phantom-head simulator. We hypothesised that early preclinical performance on a VR haptic simulator by undergraduate dental students would be associated with their preclinical and clinical performances.

MATERIALS AND METHODS

We conducted a retrospective cohort study examining practical dental performance of undergraduate dental students (2012 cohort) in the fourth year of dental school study (N=72, 46 female and 26 male) at the School of Dentistry, University of Leeds. A total of three practical test results (two preclinical and one clinical) were obtained for each student from the student education office and from the module leaders. Confidentiality was maintained by the assignment of code numbers replacing student names.

To test our hypothesis, we used two preclinical performance measures to predict clinical performance. We took the arithmetic mean of performance scores on the Simodont haptic simulator, which was used across the second year of dental study and provided a measure of haptic VR (preclinical haptic VR simulator—Y2). The Simodont haptic simulator was recently shown to have construct validity²⁸ on a range of the basic manual dexterity tasks available in the simulator courseware (ACTA, The Academic Centre for Dentistry Amsterdam, Amsterdam, Netherlands). The haptic VR trials were spread over multiple sessions as formative assessments, and the majority of runs were performed early on in the second year of dental school. To avoid the inclusion of practice trials or incomplete runs of the task, we selected only trials with a minimum task completion level of 60% and took the arithmetic mean average of all trials satisfying this threshold (the smallest number of trials for any student was 33).

Table 1 Intertest Pearson (R) correlation coefficients (n=72)

	1	2	3
Virtual reality haptic simulator	—		
Preclinical typodont crown test	−0.006	—	
Clinical crown test	0.377*	0.221	—

p < 0.001

*medium positive correlation.

For the traditional assessment, we used performance on a full crown preparation (preclinical typodont crown test—Y3) using a typodont with mounted plastic teeth on a traditional phantom-head simulator performed at year 3. Here, 40% of the score was assigned to the students' ability to critically evaluate their own performance. The outcome measure of interest was clinical performance on a full crown preparation test on a patient carried out in the fourth year of study (clinical crown test—Y4).

Data collection and statistical analysis

Preliminary analysis indicated that the scores from each variable were normally distributed (Shapiro-Wilk test, $p > 0.05$). Pearson's product-moment correlation was computed to first examine the relationship between our measures (table 1). The strength of association was interpreted based on Cohen's (1988) guidelines: small correlation ($0.1 < r < 0.3$), moderate correlation ($0.3 < r < 0.5$), and strong correlation ($r > 0.5$). Multiple regression analysis was performed to explore the relationship between students' clinical and preclinical performances, with clinical crown test performance as the dependent variable and preclinical tests (haptic VR simulator and preclinical typodont crown tests) as the predictors (independent variables). The students numerical test scores in the current study were further categorised into dichotomous (low/high performers) distinction based on each test's overall results, and the proportions of high-performing and low-performing students at each test were calculated. Fisher exact test was used to compare proportions of high-performing students at the clinical crown test with high performing students at each preclinical test. ORs and 95% CIs were calculated for high performance at the clinical crown test (dependent variable) according to high performance at the two preclinical tests (independent variable). The sensitivity and specificity of each preclinical test to predict clinical crown test performance were also calculated. The statistical significance threshold was set to $p < 0.05$. All statistical analyses were performed using IBM SPSS Statistics for Windows V.22.

RESULTS

We found no relationship between students' performance on the VR haptic simulator and preclinical typodont crown tests ($r(70) = -0.006$, $p = 0.961$; table 1). There was a weak correlation between students' performance at the clinical crown test and preclinical typodont test scores ($r(70) = 0.221$, $p = 0.062$), and there was a statistically significant medium positive relationship between students' performance on the VR haptic simulator at year 2 and the clinical crown test results at year 4 ($r(70) = 0.377$, $p = 0.001$).

To further examine the relationship between students' clinical and preclinical performances, simple linear regression analyses were conducted to examine whether overall scores on the typodont test and the VR haptic simulator could predict clinical crown test performance. We identified an independence of residuals (Durbin-Watson = 1.597) and homoscedasticity (indicated by visual inspection of a plot of studentised residuals vs unstandardised predicted values) but

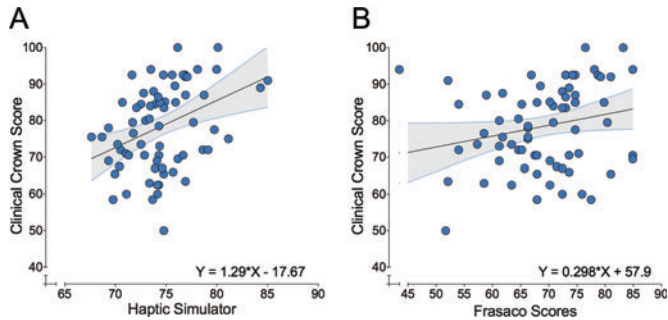


Figure 1 Regression analyses with fitted regression line and regression equations for the prediction of clinical crown test performance with (A) virtual reality haptic simulator performance as predictor and (B) preclinical typodont crown test performance as predictor. The dotted blue lines represents 95% CIs.

no evidence of multicollinearity (tolerance values >0.1). We performed robust and standard linear regressions and found minimal differences between the results and therefore report the latter here.

We found that the VR haptic simulator assessment score was a significant predictor of clinical crown performance ($F(1,70)=11.58$, $p=0.001$, $R^2=0.142$, adjusted $R^2=0.13$). The preclinical typodont crown test explained 4.9% of the clinical crown test performance with an adjusted R^2 of 3.5% ($F(1,70)=3.60$, $p=0.062$, with an R^2 of 0.049; [figure 1](#)).

ORs and 95% CIs were calculated for high performance at the clinical crown test according to high performance on the preclinical tests. This analysis indicated that students who were high performers on the VR haptic simulator assessment were 10.24 times more likely (95% CI 1.22 to 85.78) to be high performers at the clinical crown test as well (two-sided Fisher's exact $p=0.015$).

The sensitivity and specificity of the VR haptic simulator assessment and the preclinical typodont crown test to predict clinical scores were calculated. Additionally, the positive predictive value (PPV) (correctly identified high-performing students) and the negative predictive value (NPV) (correctly identified low-performing students) were also calculated. We found that the preclinical typodont crown test predicted clinical crown test performance with 67.6% sensitivity and 57.9% specificity ([table 2](#)). In contrast, the VR haptic simulator showed high sensitivity (97.1%) but low specificity (23.7%). The VR haptic simulator demonstrated high NPV (90%) compared with the preclinical typodont crown test (66.7%), with comparable PPVs (53.2%) for the former and (59%) for the latter.

Receiver operating characteristic curve analysis was performed for the two statistically significant predictors, the VR haptic

Table 2 Sensitivity, specificity, PPV, NPV and LR⁺

	Virtual reality haptic simulator	Preclinical typodont crown test
Sensitivity (%)	97.1	67.6
Specificity (%)	23.7	57.9
PPV (%)	53.2	59
NPV (%)	90	66.7
LR ⁺	1.27	1.6

LR⁺, positive likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.

Table 3 AUC_{ROC} (with 95% CI) for the preclinical predictors

	AUC	95% CI	
		Lower bound	Upper bound
Virtual reality haptic simulator assessment	0.689	0.567	0.811
Preclinical typodont crown test	0.628	0.496	0.761

AUC, area under the curve; AUC_{ROC}, Area under the receiver operating characteristic curve.

simulator and the preclinical typodont crown test. The area under the curve (AUC) for the haptic simulator was superior to that of typodont test ([table 3](#)).

Next, we explored the important observed relationship between students' mean performance across the second year of study on the VR haptic simulator and the fourth year of clinical performance in more detail. It is possible that an average measure across numerous attempts over the year may have resulted in an increase in the signal:noise ratio of this assessment, thus contributing to its superiority over the typodont (for which only one data point was available). It is also possible that the averaging of trials across the year may be masking temporal changes in the relationship between these variables. To address these possibilities, we performed two additional sets of analyses.

First, we asked whether a metric derived from a single observation of performance on the VR haptic simulator could correlate with clinical performance. To this end, we preprocessed all trial data extracted from the system (comprising 28 875 trials from 72 participants). After removing 4406 false starts (eg, student started but immediately quit the task with less than 5 s of total drilling time), 24 469 trials remained available for analysis.

From these trials, we extracted the best, worst and median trials for each participant and found that these single-trial measures did not correlate with clinical performance ($r<0.11$, $p>0.313$). However, we did note large heterogeneity in the number of attempts participants made to complete the tasks set across the module (ranging from 83 to 668 attempts). We asked whether this number, which we speculated may be indicative of participants starting from varied levels of motor skill and/or confidence, might correlate with clinical performance. We found a negative relationship between the total number of attempts and clinical performance ($r(70)=-0.297$, $p=0.011$). We then examined how many times over the course of the year participants had successfully completed tasks on the simulator. This, too, varied across participants from 11 successful trials to 293. Contrary to the idea that practice makes perfect, we found that those who had fewer attempts performed better in clinic ($r(70)=-0.276$, $p=0.019$).

Finally, we addressed the issue of temporal changes in the relationship between the VR haptic simulator and clinical performance. We separated scores from across the year into quartiles (Q1, Q2, Q3 and Q4; [figure 2](#)) to capture early and later stages of performance on the simulator for each individual participant. These measures were correlated with year 4 clinical performance, and p values were adjusted using Bonferroni correction to provide a strong control for family-wise error. The first quartile ($r(70)=0.33$, $p=0.016$) and the second quartile ($r(70)=0.32$, $p=0.021$) statistically significantly positively correlated with clinical performance, but the performance on the simulator in the last two quartiles did not (Q3: $r(70)=0.20$, $p=0.361$; Q4: $r(70)=0.03$, $p=1.0$).

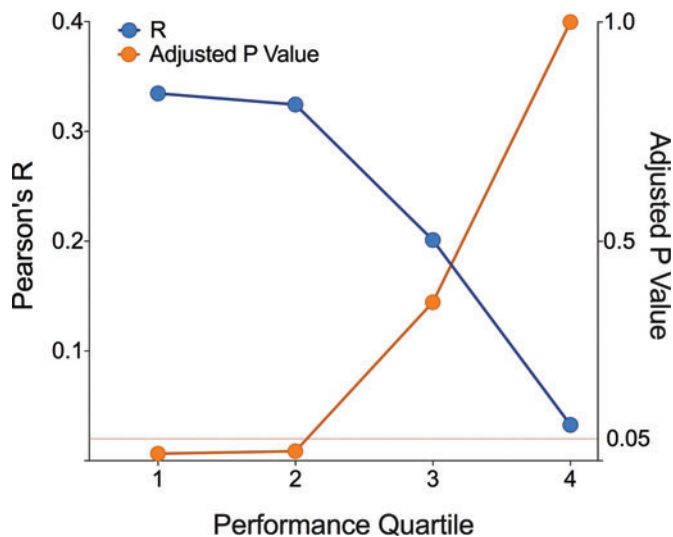


Figure 2 Time-series illustration of the relationship between performance on the virtual reality haptic simulator across year 2 and clinical crown performance in year 4. Pearson's product-moment correlation for each quartile is represented by the blue circles and plotted on the left axis. The associated Bonferroni-corrected p values are represented by orange circles and plotted on the right axis. The dotted line indicates the threshold for statistical significance.

DISCUSSION

This study shows a relationship between scores on a VR haptic simulator at the early stages of dental training and later clinical performance. Specifically, we report that 14% of the variance in clinical performance scores in year 4 of dental study could be explained by the mean performance on a variety of simulated abstract manual dexterity tasks 2 years earlier. This measure outperformed a traditional typodont test, despite the test involving the same fundamental procedure being conducted preclinically 1 year prior to the measure in clinic. Exploratory analysis of the simulator assessment indicated that the relationship with clinical performance was present only in the early stages of practice on the simulator. We consider the implications for training delivery and the strengths and limitations of these findings below.

The observation that the relationship between the VR haptic simulator and clinical performance existed only in the early stages of learning is one that requires careful examination. One perspective on these results could be that differences in starting ability at an early stage in learning may 'wash out' as students' progress through their course as all students converge on a threshold level of performance. In this way, students performing poorly at the outset catch up with the best performing students, thus weakening the relationship between the VR haptic simulator and clinical performance. It is clear that while the majority of students will reach a clinically acceptable performance threshold over the course of their studies (as was the case with the sample analysed here), it is inevitable that a small minority of individuals either will be unable or will find it extremely challenging to obtain the required levels through the standard curriculum. This is where the utility of early prediction may be at its maximal. Given that it is possible to explain a significant proportion of variance in clinical performance from performance 2 years previously, the study presents an avenue for a data-driven approach to tracking and providing timely support to individuals who may be struggling to keep up with the demands of the course.

Strengths and limitations

We examined one cohort in one dental school with access to a specific type of VR haptic simulation technology. The generalisability of these results to other cohorts, schools and simulators needs to be established as it is clear that the demonstration of clinically related objective measures of performance could have potentially important implications for the integration of VR haptic simulators in undergraduate dental training. These technologies place fewer resource demands on those delivering training relative to traditional approaches (eg, student:staff ratios, use of materials and safety). In this way, the systems used for early identification of individuals who could benefit from appropriate pedagogical support may also be valuable in delivering structured interventions to support these individuals. Recent work has shown that these tools can be effective at supporting rapid skill acquisition in novice students,³⁸ but further work is required to examine whether such approaches will prove effective for struggling students.

The context in which our data were generated must also be considered. Tasks on the VR haptic simulator were completed as part of a formative assessment, while the preclinical typodont crown test was summative. These distinct assessment approaches are likely to lead to distinctly different pressures.³⁹ One possibility may be that better performance on the formative assessments might be indicative of individual differences in motivation⁴⁰ since students were left to their own devices and could practice as much or little as they liked. We probed this hypothesis by examining the amount of times students attempted the simulated tasks over the course of the module and asked whether this correlated with clinical performance. Contrary to the motivation explanation, we found a negative correlation between attempt number and clinical performance. In other words, fewer attempts on the VR haptic simulator were linked with better scores in the clinic, thus indicating that the speed at which students are able to achieve an acceptable level of competency, even on formative assessments, is a more closely related to clinical performance than perseverance.

Finally, while only one assessment was available for the typodont, the VR haptic simulator measure comprised an average of multiple assessments over time. The increased signal:noise ratio of multiple assessments may have contributed to the increased explanatory power. Indeed, this possibility motivated our exploratory analysis, which revealed that a measure of single-trial performance, for example, the best and worst trials, showed no relationship with clinical performance. Instead, measures derived from multitrial observations held the stronger relationships with clinical performance.

CONCLUSIONS

We found performance on a VR haptic simulator at an early stage of training can predict subsequent clinical performance. This finding indicates the potential opportunities to take a data-driven approach to identifying individuals who could benefit from support in the early stages of dental training to improve their clinical performance.

Acknowledgements The authors thank the School of Dentistry Data Management team for providing access to the data reported.

Contributors LMA-S, FM, MAM-W and MM initiated the project and designed the experiment. LMA-S, supported by IM, AB, CO and AK, acquired the data. LMA-S, FM and RPM analysed the data. FM, RPM and RH worked on interpreting the empirical data. All authors provided intellectual input in drafting and revising the paper.

Funding FM, RPM and MM-W hold Alan Turing Institute Fellowships. FM and MAM-W were supported by a research grant from the Engineering and Physical Sciences Research Council (EPSRC) (EP/R031193/1).

Competing interests None declared.

Ethics approval Ethical approval to access and analyse the students' data following anonymisation (individual identifiers were replaced with unique random values) was obtained from DREC (Dental Research Ethics Committee) at the School of Dentistry, University of Leeds (DREC ref: 230915/LA/178).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs

Loulwa M Al-Saud <https://orcid.org/0000-0001-6522-8549>

Faisal Mushtaq <http://orcid.org/0000-0001-7881-1127>

Cecilie Osnes <http://orcid.org/0000-0003-4652-3854>

REFERENCES

- 1 Ranney RR, Wilson MB, Bennett RB. Evaluation of applicants to predoctoral dental education programs: review of the literature. *J Dent Educ* 2005;69:1095–106.
- 2 Urbankova A, Eber M, Engebretson SP. A complex haptic exercise to predict preclinical operative dentistry performance: a retrospective study. *J Dent Educ* 2013;77:1443–50.
- 3 Mushtaq F, Baraas RC, Al-Saud LM, et al. Should prospective dental students be screened for colour vision deficits? *Br Dent J* 2016;221:227–8.
- 4 Al-Saud LM, Mushtaq F, Mirghani Isra'a, et al. Drilling into the functional significance of stereopsis: the impact of stereoscopic information on surgical performance. *Ophthalmic Physiol Opt* 2017;37:498–506.
- 5 Mon-Williams MA, Mushtaq F, Wilkie RM, et al. A three dimensional view of stereopsis in dentistry. *Br Dent J* 2015;219:479–80.
- 6 Curtis DA, Lind SL, Plesh O, et al. Correlation of admissions criteria with academic performance in dental students. *J Dent Educ* 2007;71:1314–21.
- 7 Wood WW, Boyd MA. The DAT as a predictor of preclinical technique performance. *J Can Dent Assoc* 1982;48:599–600.
- 8 Carroll AM, Schuster GM. Correlation between students' dental admission test scores and performance on a dental school's competency exam. *J Dent Educ* 2015;79:1325–9.
- 9 Lala R, Wood D, Baker S. Validity of the UKCAT in applicant selection and predicting exam performance in UK dental students. *J Dent Educ* 2013;77:1159–70.
- 10 Poole A, Catano VM, Cunningham DP. Predicting performance in Canadian dental schools: the new CDA structured interview, a new personality assessment, and the DAT. *J Dent Educ* 2007;71:664–76.
- 11 Chamberlain TC, Catano VM, Cunningham DP. Personality as a predictor of professional behavior in dental school: comparisons with dental practitioners. *J Dent Educ* 2005;69:1222–37.
- 12 Urbankova A, Engebretson SP. The use of haptics to predict preclinic operative dentistry performance and perceptual ability. *J Dent Educ* 2011;75:1548–57.
- 13 Ballard RW, Hagan JL, Cheramie T. Relationship between hand-skill exercises and other admissions criteria and students' performance in dental school. *J Dent Educ* 2015;79:557–62.
- 14 Peterson S. The ADA chalk carving test. *J Dent Educ* 1974;38:11–15.
- 15 Walcott AM, Knight GW, Charlick RE. Waxing tests as predictors of students' performance in preclinical dentistry. *J Dent Educ* 1986;50:716–21.
- 16 Kao EC, Ngan PW, Wilson S. Wire-bending test as a predictor of preclinical performance by dental students. *Percept Mot Skills* 1990;71:667–73.
- 17 Kothe C, Hissbach J, Hampe W. Prediction of practical performance in preclinical laboratory courses - the return of wire bending for admission of dental students in Hamburg. *GMS Z Med Ausbild* 2014;31:Doc22.
- 18 Lundergan WP, Soderstrom EJ, Chambers DW. Tweezer dexterity aptitude of dental students. *J Dent Educ* 2007;71:1090–7.
- 19 Boyle AM, Santelli JC. Assessing psychomotor skills: the role of the Crawford small parts dexterity test as a screening instrument. *J Dent Educ* 1986;50:176–9.
- 20 Heintze U, Radeborg K, Bengtsson H, et al. Assessment and evaluation of individual prerequisites for dental education. *Eur J Dent Educ* 2004;8:152–60.
- 21 Causby R, Reed L, McDonnell M, et al. Use of objective psychomotor tests in health professionals. *Percept Mot Skills* 2014;118:765–804.
- 22 Boushell LW, Walter R, Phillips C. Learn-a-prep II as a predictor of psychomotor performance in a restorative dentistry course. *J Dent Educ* 2011;75:1362–9.
- 23 Zawawi KH, Afify AR, Yousef MK, et al. Reliability of didactic grades to predict practical skills in an undergraduate dental College in Saudi Arabia. *Adv Med Educ Pract* 2015;6:259–63.
- 24 Polyzois I, Claffey N, McDonald A, et al. Can evaluation of a dental procedure at the outset of learning predict later performance at the preclinical level? A pilot study. *Eur J Dent Educ* 2011;15:104–9.
- 25 Velayo BC, Stark PC, Eisen SE, et al. Using dental students' preclinical performance as an indicator of clinical success. *J Dent Educ* 2014;78:823–8.
- 26 Nunez DW, Taleghani M, Wathen WF, et al. Typodont versus live patient: predicting dental students' clinical performance. *J Dent Educ* 2012;76:407–13.
- 27 Cook DA, Brydges R, Ginsburg S, et al. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ* 2015;49:560–75.
- 28 Mirghani I, Mushtaq F, Allsop MJ, et al. Capturing differences in dental training using a virtual reality simulator. *Eur J Dent Educ* 2018;22:67–71.
- 29 Steinberg AD, Bashook PG, Drummond J, et al. Assessment of faculty perception of content validity of PerioSim, a haptic-3D virtual reality dental training simulator. *J Dent Educ* 2007;71:1574–82.
- 30 Vervoorn J, Wesselink P. The perception of the level of realism of a dental training simulator (Simodont). In: *Medicine meets virtual reality conference MMVR*, 2016.
- 31 Bakr M, Massey W, Alexander H. Students' evaluation of a 3D VR haptic device (Simodont®). does early exposure to haptic feedback during preclinical dental education enhance the development of psychomotor skills? *Int J Dent Clin* 2014;6:1–7.
- 32 Wang D, Zhang Y, Hou J, et al. IDental: a haptic-based dental simulator and its preliminary user evaluation. *IEEE Trans Haptics* 2012;5:332–43.
- 33 Al-Saud LM, Mushtaq F, Allsop MJ, et al. Feedback and motor skill acquisition using a haptic dental simulator. *Eur J Dent Educ* 2017;21:240–7.
- 34 Gray SA, Deem LP, Sisson JA, et al. The predictive utility of computer-simulated exercises for preclinical technique performance. *J Dent Educ* 2003;67:1229–33.
- 35 Imber S, Shapira G, Gordon M, et al. A virtual reality dental simulator predicts performance in an operative dentistry manikin course. *Eur J Dent Educ* 2003;7:160–3.
- 36 Urbankova A, Engebretson SP. Computer-Assisted dental simulation as a predictor of preclinical operative dentistry performance. *J Dent Educ* 2011;75:1249–55.
- 37 Polster A, Hopcraft M, Burrow M. *The application of haptic simulation in manual dexterity performance in first year dental students*. In: ConsEuro, 2015.
- 38 Al-Saud LMS, Mushtaq F, Allsop MJ, et al. Feedback and motor skill acquisition using a haptic dental simulator. *Eur J Dent Educ* 2016;11:1–8.
- 39 Rolfe I, McPherson J. Formative assessment: how am I doing? *Lancet* 1995;345:837–9.
- 40 Reeve PE, Watson CJ. An exploration of the attitudes, personality and performance of dental students. *Med Educ* 1985;19:226–37.