Measuring final-year dental students' ability to remove teeth independently using independence ratios

K-H Merbold,[1] BChD, MSc; **T C Postma,**[2] PhD

[1] Department of Maxillofacial and Oral Surgery, School of Dentistry, Faculty of Health Sciences, University of Pretoria, South Africa
[2] Department of Dental Management Sciences, School of Dentistry, Faculty of Health Sciences, University of Pretoria, South Africa

Corresponding author: T C Postma (corne.postma@up.ac.za)

Background. Universities are obliged to ensure that dental graduates possess the necessary skills to render safe and effective treatment. Empirical evidence regarding the development of safe and effective independent practice at undergraduate level is unfortunately lacking.
Objectives. To measure final-year students' abilities to correctly perform exodontia (tooth removal/extraction) skills independently, based on the applied postgraduate progressive independence theory.
Methods. Fourteen clinical teachers systematically assessed 13 263 tooth extractions completed by final-year dental students (2014 - 2016). An independence ratio (extractions performed without assistance/total number of extractions) was used as the key performance indicator to provide feedback on the ability to complete procedures independently over time. A customised index was used for controlling the level of difficulty.
Results. Final-year students (*n*=146) achieved independence ratios ranging between 90% and 94% (standard deviation 3.3%) by the end of their clinical training. Logical gradients of increased independence were illustrated with time, as well as variable performance among students. The level of difficulty index scores remained similar within cohorts per year of study. Acceptable assessment differences existed between clinical teachers, which could largely be explained by complex operational circumstances.
Conclusions. As far as we are aware, this is the first study that quantified progressive independence in exodontia for undergraduate students. The measure was sensitive enough to show logical independence gradients and variance among students. Final-year students demonstrated that they could remove >8/10 teeth independently by the time of their graduation. The measure shows promise as a proxy of competence for skills that are often practised. It is recommended that factors that influence these measurements be examined in more detail.

Afr J Health Professions Educ 2022;14(3):135-141. https://doi.org/10.7196/AJHPE.2022.v14i3.1360

Competency-based approaches in health professions education should ensure safe and effective clinical services to society.[1,2] It has been argued that outcomes in this regard should become more transparent.[2] Unfortunately, the only available empirical evidence in the surgical domain of dental education[3] is scant and unconvincing at best.

Competence is generally gauged with criterion-based methods,[4,5] which are not very feasible in busy clinical environments. Hence, workplace-based assessment (WBA) needs to be simplified to unlock true potential.[6] Giving feedback regarding a student's ability on a global scale is a useful form of big-picture task-level feedback,[7,8] but marks given to a student do not necessarily reflect their ability to perform a skill independently. We suggest that basing WBA on progressive independence[9-13] may provide a solution to such challenges owing to the relative simplicity of the measure. Progressive independence refers to a student's ability to increasingly complete procedures and skills correctly without the intervention of a clinical teacher.[9-13] Competence can be defined as the capacity to demonstrate skills independently and accurately through the application of knowledge, while displaying a professional attitude and conduct.[14-18] Students must consequently be motivated to advance their competence through increased autonomy to stimulate individual growth.[19,20] This approach does not imply that competence can be measured solely in terms of independence, because procedural, student, teacher and environmental factors may impact on independence measurement.[9] The level of difficulty is an important factor to consider and can skew independence measurement by predominantly basing independence on successful completion of easier procedures only.[9] A multifocal assessment protocol is therefore required to effectively measure independence gradients, catering for control of the level of difficulty.

Empirical evidence in this regard only exists in the field of postgraduate surgery training, which includes a supervision index[12] and the so-called Zwisch model.[21,22] The only other WBA instrument found in the undergraduate literature that would be able to measure assistance provided during clinical training, is a system proposed by Bookhan *et al.*,[23] which, to date, lacks published empirical evidence. In fact, no such evidence exists in undergraduate dental education.

Given the lack of empirical evidence of a graduate's abilities at exit level in undergraduate dental education, we aimed to measure students' abilities to remove teeth independently in the final year of study in a WBA environment.

Materials and methods
Institutional context
This study was conducted in the exodontia (tooth removal/extraction) clinic of the Department of Maxillofacial and Oral Surgery, School of Dentistry, University of Pretoria (UP), South Africa (SA), in 13 dental chairs in open-plan cubicles, allowing for unrestricted supervision of multiple students. After each encounter, students were directly assessed on chairside computers using the password-protected GoodX Dental Studio software (GoodX, South Africa). Authorised clinical teachers (henceforth referred to

as teachers) could sign off assessments on the system, with encrypted data being stored on a central server. Each day comprised three 2-hour sessions, with a student typically attending to 1 or 2 patients per session. The student-to-teacher ratio was ~4:1. Thirteen trained teachers assisted the primary researcher throughout the study period (2014 - 2016).

Study design and sample selection

The study cohort comprised all enrolled final-year (fifth year) dental students (2014 - 2016). Their first exposure to removal of teeth was at the beginning of their fourth year of study.

Measurements

After a thorough case presentation, including radiographic examination, students and teachers agreed on the difficulty (easy, standard or difficult) of the intended extraction, based on specific criteria (Table 1).

After completion, students captured the dental procedure codes (8201/8202 (first/subsequent extraction per quadrant), 8937 (surgical extraction), 8941 (first impaction), 8943 (second impaction), 8945 (subsequent impaction)). Teachers rated the outcome of each extraction based on a scale (Table 1). A zero score applied to demonstrations to inexperienced students, typically at the beginning of their fourth year or at first exposures to surgical extractions. Scores of 1 - 3 applied to students requiring teacher intervention (independence score 0), taking difficulty level into account. Similarly, scores of 4 - 6 applied to cases where students removed teeth independently (independence score 1), sometimes with verbal advice. Teachers could upgrade difficulty to the next level (from easy to standard and standard to difficult) after completion. This step was warranted if unforeseen complications arose related to tooth/bone anatomy (not visible clinically or on two-dimensional radiographs) or because of patient behaviour (e.g. children <12 years old).

Relevant data were obtained from an electronic Microsoft Office Excel (Microsoft Corp., USA) report, downloaded from the GoodX Dental Studio database, which was used to calculate the number of teeth removed (including teeth removed surgically), independence ratios and difficulty scores. These indicators were reported to students as feedback on a quarterly basis and at the end of the cycle to encourage them to be more engaged regarding the procedure, attempt more difficult procedures and increase their independence.

The independence ratio over time and a level of difficulty index (LDI) were measured as follows: as A, B and C (Box 1) had varying levels of difficulty, a utility weight had to be assigned to each. It was decided to use rounded-off monetary values (based on relative value units) assigned to dental procedure codes used by Discovery Health, the largest open medical aid scheme in SA (https://www.discovery.co.za/portal/individual/login – password-protected information available from the researcher). The removal of a second tooth (code 8202) in the same region as a first extraction is generally easier than the first (code 8201); hence, a lower monetary value applies. The medical aid rate for a conventional extraction was 2.5 times more than for the removal of a second or subsequent tooth. Similarly, surgical removal of a tooth (code 8937) was ~9 times more expensive than the cost of removing an additional tooth. Utility weightings of 1, 2.5 and 9 were therefore respectively allocated to the easy (A), standard (B) and difficult (C) extractions. It should be noted that the aim of this index was merely to serve as a proxy of difficulty to ensure valid interpretation of independence ratios and as feedback to students on difficulty achieved. Data modelling shows that a student with a higher number of easy extractions will achieve a lower difficulty index than a student with fewer easy extractions, when the number of standard and difficult extractions are the same for both students. The reverse is also true when a student achieves more difficult extractions while the number of easy and standard extractions is the same.

During pretesting of the data, tooth extraction (removal) counts displayed skewed distributions, while the LDI and independence ratios (Box 1) showed close to normal distributions. Changes in extraction counts, LDI scores and independence ratios were reported by using descriptive statistics. The Wilcoxon signed-rank test, paired samples t-tests and χ^2 tests were used for the three indicators, respectively, to measure changes in progression between quarterly feedback episodes.

The primary researcher performed the same role as all the other teachers and his data were used as reference to evaluate data variance of the others because of his leading and staff developmental role. For independence ratio differences, χ^2 analysis was used to assess these differences. The teachers' average difficulty scores (4 - 6) for successfully completed conventional tooth extractions were compared with the primary researcher's average ratings at year-end. Surgical tooth extractions were excluded because of their high level of difficulty (default score of 6) and as some teachers had to supervise more surgical tooth extractions than others owing to the operational environment. A 10% difference was arbitrarily set to be acceptable owing to expected variances in work allocation and supervision roles due to operational circumstances and other environmental factors.[9]

Table 1. Continuous assessment scales for exodontia (tooth removal/extraction)

Level of difficulty		Independence scale	
Score	Criteria*	Score	Criteria
0	Easy/standard/difficult	Not assessed	Completed by the clinical teacher (e.g. demonstration)
1	Easy	0	Completed by the student, with physical intervention of the clinical teacher
2	Standard		
3	Difficult		
4	Easy	1	Completed by the student, without physical intervention of the clinical teacher. Oral advice may however have been provided
5	Standard		
6	Difficult		

*Easy = periodontally compromised/mobile teeth, primary tooth with root resorption, bone loss around roots, e.g. periapical abscess or periapical bone destruction visible on a radiograph; standard = normal extraction of primary or permanent teeth with normal bone and roots; difficult = dilacerated or long or divergent roots, root canal-treated tooth without bone loss, root rest with a difficult approach, impactions.

Box 1. Independence ratio and level of difficulty formulas

$$\text{Independence ratio} = \frac{\text{Number of times the skill was completed correctly and independently}}{\text{Number of times the skill was completed}}$$

$$\text{Level of difficulty index} = \frac{(1 \times A) + (2.5 \times B) + (9 \times C)}{(A + B + C)}$$

A = teeth for which a score of 4 was obtained, n
B = teeth for which a score of 5 was obtained, n
C = teeth for which a score of 6 was obtained, n

Some of these teachers were postgraduate students specialising in surgery. By default, they had better surgical skills and were often approached to assist students when complications were anticipated, which may have skewed the LDI to be higher than normal.

Students' exposure to teachers was also evaluated, as it was anticipated through the literature[9] that in ensuring patient safety, some teachers have more courage to allow independence than others. It was argued that this could not be controlled through precalibration and that it would be acceptable provided students were exposed to most of the teachers.

Ethical approval

The Research Ethics Committee, Faculty of Health Sciences, UP, gave ethical approval (ref. nos 189/2015 and 137/2016) to collect the data prospectively for the 2015 and 2016 cohorts and retrospectively for the 2014 cohort.

Results

A total of 13 263 assessments (2014 - 216) were included after discarding three obvious incorrect mark allocations and 166 extractions that were not assessed.

The 2014, 2015 and 2016 cohorts' mean exposure to the range of teachers amounted to 88% (range 60 - 100%), 87% (73 - 100%) and 90% (70 - 100%), respectively.

Student statistics

Fig. 1 provides a visual representation of the gradient of the average independence growth within cohorts between feedback episodes, clearly flattening off towards the end.

Independence ratios started at 0.86 (standard deviation (SD) 0.09; range 0.56 - 1.00) in 2014, increasing to 0.90 (0.05; 0.79 - 1.00) by the end of training (Table 2). In 2015, the baseline was 0.89 (0.10; 0.53 - 1.00), increasing to 0.92 (0.04; 0.78 - 1.00), and in 2016 the baseline was 0.92 (0.06; 0.79 - 1.00), progressing to 0.94 (0.03; 0.86 -1.00) (Table 2). As tooth extraction counts increased, only slight (often statistically insignificant) variable changes could be detected for the mean LDI scores, as measured during feedback episodes (Table 2). LDI scores varied considerably at baseline among students. The variability declined somewhat towards the end of the year, with smaller SDs being observed.

Table 3 displays examples of feedback provided to individual students (2014 - 2016), showing achievements in the number of teeth removed, LDI scores and independence ratios. Students were provided with an anonymised table quarterly, similar to Table 3, informing them of:
- problem areas that need attention
- progression/worsening compared with the previous feedback episode
- standing in relation to their peers.

Values in bold in Table 3 show performance below par for extraction count, LDI and independence ratios at first and last feedback for 6 students chosen for illustrative purposes. The extraction count threshold was 60 (typical minimum extraction quota for the year), while the LDI threshold was 2.35 (minimum requirement of achieving 90% standard extractions and 10% easy extractions). The independence ratio threshold was 90% (arbitrary chosen target for students). Addendum A (https://www.samedical.org/file/1880) contains a similar display for the entire study. The number of students not complying with the independence ratio target decreased from first to last feedback in each of the cohorts. Moreover, most of those who did not comply initially, highlighted at the bottom of Addendum A, appear
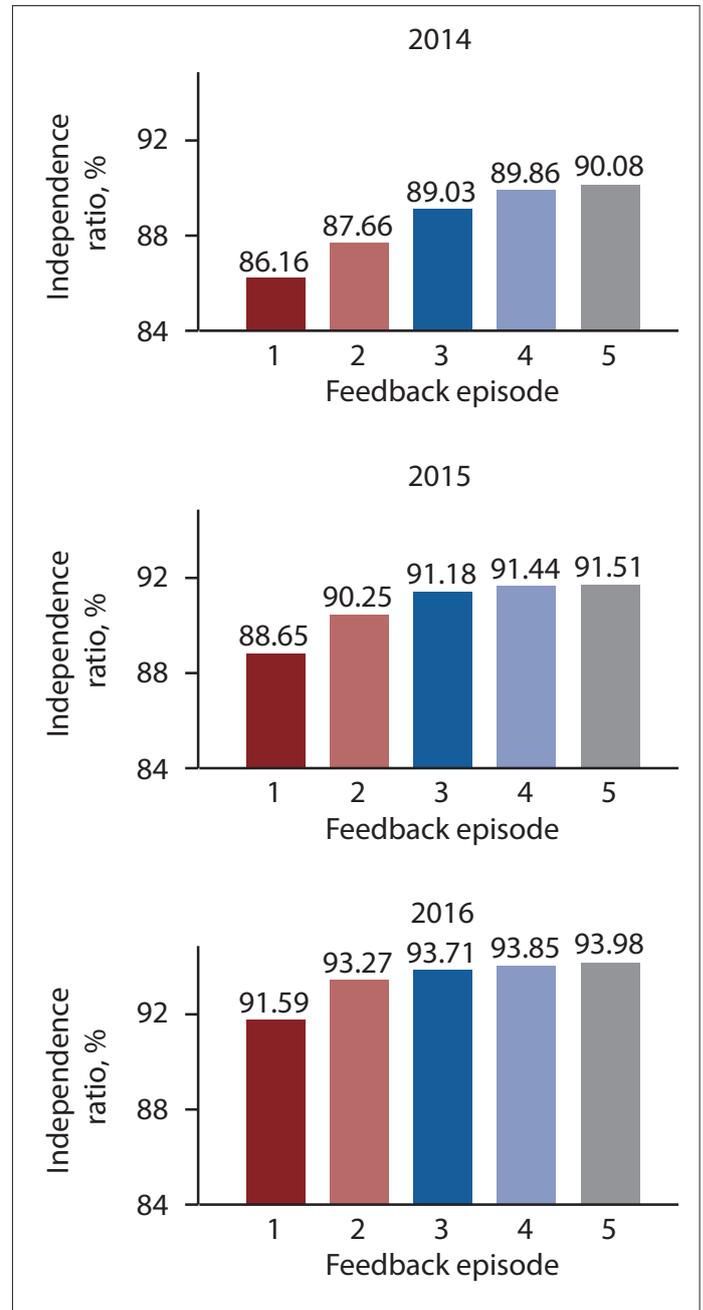

Fig. 1. Average growth in independence (2014 - 2016).

to have improved their independence ratios over time (some more than others), with very few ending below 85%. It should, however, be noted that 3/29, 4/25 and 13/45 of those who achieved the 90% independence ratio threshold in 2014, 2015 and 2016, respectively, did not conform to the LDI standard of 2.35. Conversely, many students with slightly below threshold independence ratios achieved LDI scores above the threshold, with only a minority achieving below par independence ratios and LDI scores (Addendum A).

Clinical teacher statistics

Table 4 shows the clinical teacher statistics and their years of experience. Independence ratios awarded by teachers were similar to those awarded

Table 2. Changes in tooth extraction counts, LDI scores and independence ratios (2014 - 2016)

Cohort	Description	March/April		June		August/September		September		October
		1st	p	2nd	p	3rd	p	4th	p	5th
014	Extraction count*		0.000		0.000		0.000		0.000	
	Median	33		55		73		84		87
	Range	13 - 75		17 - 121		25 - 153		26 - 162		27 - 168
	LDI		0.405		0.072		0.059		0.286	
	n	2 060		3 326		4 367		4 965		5 207
	Mean	3.06		3.02		2.97		2.94		2.93
	SD	0.71		0.59		0.55		0.50		0.49
	SE	0.09		0.08		0.07		0.07		0.06
	Range	2.03 - 4.55		2.11 - 4.41		2.05 - 4.58		2.11 - 4.50		2.20 - 4.43
	Independence ratios		0.026		0.000		0.000		0.042	
	Mean	0.86		0.88		0.89		0.90		0.90
	SD	0.09		0.07		0.06		0.05		0.05
	SE	0.01		0.01		0.01		0.01		0.01
	Range	0.56 -1.00		0.74 - 1.00		0.75 - 1.00		0.78 - 1.00		0.79 - 1.00
2015	Extraction count*		0.000		0.000		0.000		0.000	
	Median	30		57		78		89		92
	Range	17 - 66		39 - 147		354 - 190		64 - 207		65 - 219
	LDI		0.886		0.066		0.327		0.328	
	n	1 143		2 230		2 996		3 399		3 477
	Mean	2.75		2.74		2.68		2.69		2.7
	SD	0.74		0.59		0.6		0.59		0.58
	SE	0.12		0.1		0.1		0.1		0.1
	Range	1.47 - 4.62		1.80 - 4.50		1.84 - 4.50		1.76 - 4.67		1.77 - 4.73
	Independence ratios		0.081		0.013		0.123		0.283	
	Mean	0.89		0.9		0.91		0.91		0.92
	SD	0.1		0.06		0.05		0.04		0.04
	SE	0.02		0.01		0.01		0.01		0.01
	Range	0.53 - 1.00		0.72 - 0.98		0.74 - 0.98		0.77 - 0.98		0.78 - 0.98
2016	Extraction count*		0.000		0.000		0.000		0.000	
	Median	23		56		77		83		85
	Range	9 - 62		29 - 115		50 - 127		58 - 148		65 - 150
	LDI		0.442		0.917		0.427		0.151	
	n	1 295		3 071		4 072		4 440		4 579
	Mean	2.66		2.74		2.74		2.73		2.71
	SD	0.86		0.69		0.58		0.57		0.56
	SE	0.12		0.1		0.08		0.08		0.08
	Range	1.30 - 4.67		1.49 - 4.61		1.61 - 4.45		1.61 - 4.31		1.61 - 4.43
	Independence ratios		0.015		0.104		0.406		0.053	
	Mean	0.92		0.93		0.94		0.94		0.94
	SD	0.06		0.04		0.04		0.03		0.03
	SE	0.01		0.01		0.01		0		0
	Range	0.79 - 1.00		0.83 - 1.00		0.82 - 1.00		0.86 - 1.00		0.86 - 1.00

LDI = level of difficulty index; SD = standard deviation; SE - standard error.
*Teeth removed by the student, n (statistical tests: Wilcoxon signed-rank test, LDI - paired samples t-test, independence ratios – χ2 test).

by the primary researcher (the reference), with a few exceptions. In 2014, teachers 1 and 8 (*n*=2/9 teachers) awarded 3% and 4% higher independence (χ^2; $p<0.05$), respectively, compared with the reference. Their mean difficulty scores were, however, similar to those of the primary researcher. The same occurred in 2015 (teachers 2 and 3; *n*=2/11). In 2016, only teacher 9 was 5% less lenient in allowing independence.

Average difficulty rating differences compared with those of the primary researcher mostly fell within the predetermined limit of 5% (Table 4). In 2014,

8/9 teachers differed by <5%. The highest observed difference in 2014 was 7.6% for teacher 7, who had a very small extraction assessment count (*n*=49). In 2015, only 2/11 average difficulty ratings exceeded the 5% cut-off slightly (5.91% and ⁻5.70%). In 2016, 2/9 differences of 5% were noted, specifically for teachers 10 (14.9%) and 12 (7.1%) (Table 4).

Discussion

As far as we are aware, this is the first study that illustrates the development

Table 3. First and last feedback provided to 6 students on tooth extraction counts, level of difficulty index scores and independence ratios*

Students, n	First feedback				Last feedback			
	Extraction count	LDI	Independence ratio	Targets achieved	Extraction count	LDI	Independence ratio	Targets achieved
3	75	2.38	0.99	All	123	2.79	0.98	All
16	25	2.34	0.88	None	51	2.31	0.92	For independence ratio only
17	38	1.47	0.92	For independence ratio only	82	2.26	0.93	For extraction count and independence ratio only
18	30	3.37	0.87	For LDI only	76	3.14	0.92	All
41	26	2.03	0.69	None	97	2.22	0.88	For extraction count only
28	33	2.27	0.85	None	91	2.66	0.90	All

LDI = level of difficulty index.
*Values in bold = performance below par for extraction count, LDI and independence ratios.

Table 4. Comparison among clinical teachers' assessment load, independence ratios and level of difficulty

Teacher	Reference	1	2	3	4	5	6	7	8	9	10	11	12	13
Experience, years	>20	>20	>20	>20	>10	>20	<5	>20	<5	<10	<10	>20	>10	<5
2014														
n	729	297	866	633	685	688	239	49	433	235	-	-	-	-
Independence ratio, mean	0.94	0.97*	0.96	0.96	0.95	0.93	0.91	0.92	0.98*	0.94	-	-	-	-
Difficulty, mean	4.83	4.49	4.73	4.85	4.96	5.06	4.93	5.20	4.76	4.92	-	-	-	-
SD	0.57	0.69	0.55	0.56	0.41	0.24	0.67	0.59	0.66	0.64	-	-	-	-
95% CIL	4.79	4.86	4.69	4.81	4.93	5.04	4.84	5.02	4.70	4.84	-	-	-	-
95% CIU	4.88	5.02	4.77	4.90	4.99	5.08	5.02	5.38	4.83	5.01	-	-	-	-
Difference to the reference, %	-	2.28	−2.07	0.41	2.69	4.76	2.07	7.66	−1.45	1.86	-	-	-	-
2015														
n	260	199	533	414	331	332	13	187	260	159	-	501	-	-
Independence ratio, mean	0.93	0.96	0.98*	0.97*	0.96	0.93	0.85	0.96	0.94	0.93	-	0.94	-	-
Difficulty, mean	4.74	4.86	4.70	4.61	4.86	5.02	4.91	4.83	4.62	4.91	-	4.47	-	-
SD	0.59	0.75	0.60	0.56	0.52	0.40	0.54	0.51	0.77	0.61	-	0.56	-	-
95% CIL	4.67	4.75	4.65	4.55	4.81	4.97	4.55	4.75	4.53	4.81	-	4.42	-	-
95% CIU	4.82	4.97	4.75	4.66	4.92	5.06	5.27	4.90	4.72	5.01	-	4.52	-	-
Difference to the reference, %	-	-	−0.84	−2.90	2.53	5.91	3.59	1.90	−2.53	3.59	-	−5.70	-	-
2016														
n	245	293	866	633	-	-	-	-	280	390	192	662	165	565
Independence ratio, mean	0.97	0.98	0.98	0.94	-	-	-	-	0.96	0.92*	0.97	0.98	0.96	0.98
Difficulty, mean	4.67	4.90	4.53	4.66	-	-	-	-	4.58	4.88	5.37	4.44	5.00	4.67
SD	0.63	0.67	0.61	0.59	-	-	-	-	0.69	0.58	0.50	0.58	0.63	0.63
95% CIL	4.59	4.82	4.49	4.61	-	-	-	-	4.50	4.82	5.29	4.40	4.90	4.58
95% CIU	4.75	4.98	4.57	4.71	-	-	-	-	4.66	4.94	5.44	4.49	5.10	4.69
Difference to the reference, %	-	4.93	−3.00	−0.21	-	-	-	-	−2.00	4.50	14.99	−4.93	7.07	−0.07

SD = standard deviation; CIL = confidence interval lower; CIU = confidence interval upper.
*χ^2: $p<0.05$.

of independence involving undergraduate students. The design of the WBA instrument comprised three inter-related elements, i.e. independence ratios, a customised LDI and tooth extraction counts. Each of these performance indicators plays an important role in motivating students to develop their skills.

Variable independence ratios observed at baseline suggest different abilities at the start, ranging from 53% to 100%. As the extraction count increased, independence improved, but not all students could reach independence of 80% (especially in 2014 and 2015). The general trend, however, was a gradient of increased independence.

The validity of independence ratios is highly dependent on the control of difficulty to avoid the risk of manipulation of the ratios by only attempting easy procedures, which by default results in high ratios. This study showed that, when measured over long periods of time, average difficulty levels remained approximately the same, which is logical, as work is randomly allocated. LDI scores varied considerably at baseline, suggesting that some students were already engaging in more complex procedures than others. Although the mean LDI remained relatively consistent, the SDs shrunk when compared with baseline, suggesting that lower-performing students

started engaging in more complex procedures. Unfortunately, ~20% of students who achieved the 90% independence ratio threshold did not conform to the LDI threshold of 2.35, suggesting inflated independence ratios and low proficiency and/or engagement in complex work. Conversely, most students who achieved below par independence ratios of 80 - 89%, had final LDI scores above the threshold. The latter scenario can be deemed more acceptable than the previous one, because these students achieved some success in independently completing more difficult work.

Variances in clinical teacher statistics suggest that teacher and environmental factors may have affected the scoring process. The literature indicates that teacher audacity to trust a student to continue ensuring patient safety,[24] is indeed a critical factor in progressive independence.[9] Clearly, there were teachers who allowed slightly less independence in the current study, but it could not be linked to years of experience (Table 4). It can be argued that environmental factors[9] such as the scheduling of teachers may also influence the results. Teachers assessing only in the latter half of the year by default award higher independence ratios than those working throughout the year, as students' competence increases with time. This is evident in the primary researcher's slightly higher than normal independence ratios awarded for 2016 (Table 4). He worked in the clinic only in the latter half of the year (owing to a medical problem), when the students had more confidence. Further study is required to understand the relationships of interacting forces related to the students, teachers, environment, procedures and feedback, specifically to inform staff development.

Teachers' LDIs were more reflective of the complex environment. Only broad congruence could be shown, with some variability in interpretation. The particularly large deviation of teacher 10 in 2016 can be explained by his role in supervising cases that were anticipated in advance to be problematic (when the primary researcher was absent). These cases included, e.g. root canal-treated teeth, teeth with divergent roots and extreme bone density. The remaining deviations of ~5% can be explained by environmental factors in work allocation and the interpretation of what is considered to be difficult, in particular. Such interpretation can only be calibrated up to a certain point and in a real-life situation. A recent article suggests that teacher variance should be supported and an assortment of teachers who manage students be encouraged.[25] Students were exposed to a variety of teachers. The results of the study show that the majority of students were indeed exposed to most teachers, which in turn would even out discrepancy to some degree. Given all these interacting factors, we were satisfied that the LDI gave a fair reflection of the difficulty achieved. It is however recognised that continuous staff training and communication in this regard are essential.

Study limitations
The data reported in this study reflect the nature of operational circumstances, which affect WBA, but are deemed reliable enough to be used and to be developed further. A clearer understanding of interacting factors is needed to improve the system to further enhance reliability and validity.

The effect of feedback was not examined. The progression shown between feedback episodes cannot be attributed to feedback, because of the cross-sectional nature of the snapshots taken to periodically measure independence. The association between progression and feedback should therefore be explored in further studies.

This pilot study involved measurement in final-year cohorts only and concentrated on a single range of skills. Applying the system to a broader range of scenarios is essential to evaluate its potential. Expansion to other disciplines (e.g. radiography and restorative dentistry), where high numbers of procedures are performed, is feasible.

Standards used in this study were also not used to pass and fail students, because of the development of the new WBA system as part of a research project. For this period, other assessments were used for final pass/fail decisions. WBA was used as a formative assessment tool based on data, encouraging students to aim for targets. If distinct cut-offs existed, student behaviour may also have differed. These issues also need further exploration.

It is indeed so that no educational intervention is perfect owing to considerable complexity of the context.[26] Attempts to apply theory in context with the aim of improvement and to generate new theory are however essential,[26] which is the aim of this ongoing project.

Conclusion
This study illustrates that it is possible to achieve independence in exodontia, ranging between 80% and 100%, at the exit-level of a dental curriculum, and is the first known empirical evidence in this regard. Gradients of increased independence could be shown over time, as well as variance between students. Independence ratios appeared to be sensitive enough to discriminate between students, with high and low dependence. Independence ratios should be used in combination with difficulty scores to draw inferences on ability. The measurement of difficulty, student exposure to different teachers and a relatively high number of procedures are, however, critical to increase reliability in the complex environment.

1. Lockyer J, Carraccio C, Chan M-K, et al. Core principles of assessment in competency-based medical education. Med Teach 2017;39(6):609-616. https://doi.org/10.1080/0142159X.2017.1315082
2. Moore U, Durham J. Invited commentary: Issues with assessing competence in undergraduate dental education. Eur J Dent Educ 2011;15(1):53-57. https://doi.org/10.1111/j.1600-0579.2010.00634.x
3. Durham JA, Moore UJ, Corbett IP, Thomson PJ. Assessing competency in dentoalveolar surgery: A 3-year study of cumulative experience in the undergraduate curriculum. Eur J Dent Educ 2007;11(4):200-207. https://doi.org/10.1111/j.1600-0579.2007.00455.x
4. Sadler DR. Interpretations of criteria-based assessment and grading in higher education. Assess Eval High Educ 2005;30(2):175-194. https://doi.org/10.1080/ 0260293042000264262
5. Wagaarachchi PT, Graham WJ, Penney G, et al. Holding up a mirror: Changing obstetric practice through criterion-based clinical audit in developing countries. Int J Gynecol Obstet 2001;74(2):119-130. https://doi.org/10.1016/S0020-7292(01)00427-1
6. Norcini J, Burch V. Workplace-based assessment as an educational tool. AMEE Guide No. 31. Med Teach 2009;29(9-10):855-871. https://doi.org/10.1080/014215907017 75453
7. Doyle JD, Webber EM, Sidhu RS. A universal global rating scale for the evaluation of technical skills in the operating room. Am J Surg 2007;193(5):551-555. https://doi.org /10.1016/j.amjsurg.2007.02.003
8. Hattie J, Timperley H. The power of feedback. Rev Educ Res 2007;77(1):181-182. https://doi.org/10.3102/003465430298487
9. Dijksterhuis MG, Voorhuis M, Teunissen PW, et al. Assessment of competence and progressive independence in postgraduate clinical training. Med Educ 2009;43(12):1156-1165. https://doi.org/10.1111/j.1365-2923.2009.03509.x
10. Franzone JM, Kennedy BC, Merritt H, et al. Progressive independence in clinical training: Perspectives of a national, multispecialty panel of residents and fellows. J Grad Med Educ 2015;7(4):700-704. https://doi.org/10.4300/JGME-07-04-51
11. Kashner TM, Byrne JM, Henley SS, et al. Measuring progressive independence with the resident supervision index: Theoretical approach. J Grad Med Educ 2010;2(1):8-16. https://doi.org/10.4300/JGME-D-09-00083.1

12. Kashner TM, Byrne JM, Chang BK, et al. Measuring progressive independence with the resident supervision index: Empirical approach. J Grad Med Educ 2010;2(1):17-30. https://doi.org/10.4300/1949-8357-2.1.17
13. Kennedy TJ, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: A tradition worth defending? Acad Med 2005;80(10):S106-S111. https://doi. org/10.4300/1949-8357-2.1.17
14. Khan K, Ramachandran S. Conceptual framework for performance assessment: Competency, competence and performance in the context of assessments in healthcare – deciphering the terminology. Med Teach 2012;34(11):920-928. https://doi.org/10.3109/0142159X.2012.722707
15. Chuenjitwongsa S, Oliver RG, Bullock AD. Competence, competency-based education, and undergraduate dental education: A discussion paper. Eur J Dent Educ 2018;22(1):1-8. https://doi.org/10.1111/eje.12213
16. Evans A. Assessing competence in surgical dentistry. BDJ 2001;190(7):343-346.
17. General Dental Council. The First Five Years. A Framework for Undergraduate Dental Education. London: GDC, 2002.
18. General Dental Council. The First Five Years (Interim). London: GDC, 2008.
19. Schumacher DJ, Bria C, Frohna JG. The quest toward unsupervised practice: Promoting autonomy, not independence. JAMA 2013;310(24):2613-2614. https://doi.org/10.1001/ jama.2013.282324
20. Edwards PC. The future of dental education: Toward disruptive innovation or incremental improvements? Oral Surg Oral Med Oral Pathol Oral Radiol 2015;119(3):257-259. https://doi.org/10.1016/j.oooo.2014.12.006
21. DaRosa DA, Zwischenberger JB, Meyerson SL, et al. A theory-based model for teaching and assessing residents in the operating room. J Surg Educ 2013;70(1):24-30. https://doi.org/10.1016/j.jsurg.2012.07.007
22. George BC, Teitelbaum EN, Meyerson SL, et al. Reliability, validity, and feasibility of the Zwisch scale for the assessment of intraoperative performance. J Surg Educ 2014;71(6):e90-e96. https://doi.org/10.1016/j.jsurg.2014.06.018
23. Bookhan V, Becker LH, Oosthuizen MP. Criteria referenced student self-assesment in restorative dentistry. S Afr Dent J 2005;60(4):161-166.
24. Ten Cate TJO, Snell L, Carraccio C. Medical competence: The interplay between individual ability and the health care environment. Med Teach 2010;32(8):669-675. https://doi.org/10.3109/0142159X.2010.500897
25. Boursicot K, Kemp S, Wilkinson T, et al. Performance assessment: Consensus statement and recommendations from the 2020 Ottawa Conference. Med Teach 2020;14:1-10. https://doi.org/10.1080/0142159X.2020.1830052
26. McKenney S, Reeves TC. Conducting Educational Design Research. 2nd ed. London: Routledge, 2018.