

Workplace-based assessment of final year students in Paediatric Dentistry at the University of the Western Cape: Is it an indication of clinical competence ?

by
Nadia Mohamed



Research Assignment submitted in partial fulfilment of the requirements for the degree of Masters of Philosophy in Health Professions Education at the Faculty of Medicine and Health Sciences, Stellenbosch University

iyUNIVESITHI
STELLENBOSCH
UNIVERSITY



Supervisor: Dr. E.J. Smit
Co-supervisor: Prof. S.C. Van Schalkwyk
Stellenbosch University

December 2018

DECLARATION

I, the undersigned, hereby declare that the work contained in this assignment is my original work and that I have not previously submitted it, in its entirety or in part, at any university for a degree.

Date: December 2018

Signature:

Copyright © 2018 Stellenbosch University

All rights reserved

ACKNOWLEDGEMENTS

I would like to thank the following individuals for their role in helping me complete this thesis:

- ◆ My supervisor, Dr Smit, for her support during this entire process- Your encouragement, support and calm demeanour made it easier for me to manage the pressure, especially towards the end. Research in Health Professions Education was entirely new to me and I learnt a tremendous amount. Thank you for your patience.
- ◆ My co-supervisor, Prof Van Schalkwyk, for her insight and guidance, especially in the refinement stage of this thesis.
- ◆ Dr Carl Lombard for the statistical analysis.
- ◆ My best friend, Craig- this journey would not have been the same without you.

Thank you!

This thesis is dedicated to my students- for challenging me to find ways to improve my teaching practices and for being the driving force behind pursuing a degree in Health Professions Education.

TABLE OF CONTENTS

	Page number
List of appendices	v
List of figures	vi
List of tables	vii
List of abbreviations	viii
Summary	1
Opsomming	3
Chapter 1: Introduction	5
Chapter 2: Literature review	8
2.1 Purpose of assessment	8
2.2 Principles of sound assessment	9
2.3 Assessment of clinical competence	15
Chapter 3: Current assessment practices for the final year Paediatric Dentistry module	20
3.1 Assessment Plan	20
3.2 Implementation of continuous assessment tasks	23
3.2.1: Mini-CEX evaluations and logbook quota	23
3.2.2: Case presentations	26
3.2.3: Blinded patient case	26
Chapter 4: Research design and methodology	28
4.1 Problem statement	28
4.2 Research question	28
4.3 Aim and objectives	28

Table of contents (...cont)		Page number
4.4	Study design	28
4.5	Setting	29
4.6	Target and study population	29
4.7	Data collection and analysis	29
4.8	Ethical considerations	31
4.9	Positioning of the researcher	32
Chapter 5: Results		33
5.1	Summary of assessment results	33
5.2	Correlation between the various components of the continuous WPBA	34
5.3	Identification of struggling students	38
5.4	Reliability of WPBA	39
5.5	Validity of WBPA	43
Chapter 6: Discussion		44
6.1	The impact of continuous formative assessment	44
6.2	Formative results as predictors for summative performance	45
6.3	The value of multiple and varied assessment	46
6.4	The case for reliability	47
6.5	A case for validity	50
6.6	Limitations of the study	51
Chapter 7: Conclusion & Recommendations		53
References		54

LIST OF APPENDICES

Appendix	Description	Page number
A	Minimal procedural requirements/ quota	62
B	5 th year module descriptor	63
C	Current clinical rubric used in Paediatric Dentistry (holistic)	64
D	Excerpt of detailed module descriptor used to compile clinical rubric	66
E	Analytic rubric used for blinded patient case	67
F	Ethics approval: University of Stellenbosch	68
G	Ethics approval: University of the Western Cape	69

LIST OF FIGURES

Figure number	Description	Page number
Fig. 2.3.1	Adaptation of Miller's pyramid	16
Fig. 3.1.1	Overall assessment plan (Paediatric Dentistry)	20
Fig. 3.1.2	Clinical assessment components (Paediatric Dentistry)	22
Fig. 3.1.3	Weighting of assessment components	22
Fig. 5.1.1.	Summary of assessment data	34
Fig. 5.2.1	Scatterplot matrix depicting correlation between the various assessment components for the combined 2016-2017 cohort	34
Fig. 5.2.2	Correlation between average mini-clinical evaluation and the BPC	35
Fig. 5.2.3	Correlation between the case presentations (average) and the BPC	36
Fig. 5.2.4	Correlation between average mini-clinical evaluation and percentage of logbook quotas completed	36
Fig. 5.2.5	Correlation between the combined continuous WPBA and BPC	37
Fig. 5.4.1.1	Allocated scores for the mini-CEX encounters over time (2016)	39
Fig. 5.4.1.2	Allocated scores for the mini-CEX encounters over time (2017)	40

LIST OF TABLES

Table number	Description	Page number
Table 5.1.1	Summary of assessment data for the cohort	33
Table 5.2.1	Spearman correlation coefficients between various assessment components	35
Table 5.3.1	Summary of marks of students who failed a component of the assessment tasks	38
Table 5.4.1.1	Scoring profiles of clinical supervisors based on scores given according to rubric (scores 1 to 5) for diagnostic encounters	41
Table 5.4.1.2	Interpretation of a score of 4	41
Table 5.4.2.1	Evaluation of reliability of the assessment approach using qualitative strategies	42
Table 5.5.1	Validity argument for WPBA in Paediatric Dentistry	43

LIST OF ABBREVIATIONS

BPC	Blinded Patient Case
BPE	Blinded Patient Encounters
CbD	Case-based Discussions
CEC	Clinical Encounter Cards
DOPS	Direct Observation of Procedural Skills
GCP	Good Clinical Practice
Mini-CEX	Mini-Clinical Evaluation Exercise
r_s	correlation
UWC	University of the Western Cape
WPBA	Workplace-based assessment

SUMMARY

Clinical competence is multifaceted and requires the integration of knowledge, skills and attitudes. The clinical environment where patients are treated, provides an opportunity for student assessment of clinical competence in an authentic workplace setting at the 'does' level of behaviour.

Final year dental students in the Department of Paediatric Dentistry at the University of the Western Cape are assessed in the clinical environment on a daily basis through multiple clinical evaluations over the course of a year. An additional end-of-module clinical assessment in the form of a single blinded patient case (BPC) is required to decide if students have reached the expected level of clinical competence. Both the reliability and feasibility of this single end-of-module clinical case have been questioned in this setting. The utility of continuous formative WPBA during a course in determining progression gained at the end of a programme is however still being debated in the literature.

This study aimed to determine if the current continuous WPBA results in the Department of Paediatric Dentistry could be used as an indication of clinical competence of final year students at the end of the module. A retrospective, quantitative, cross-sectional study was conducted of all complete assessment records of final year students (2016- 2017) from the Paediatric Dentistry Department at UWC. The characteristics of, and correlation between, the continuous WPBA components were analysed together with an evaluation of the reliability and validity of the assessment results.

On average, students achieved the highest score for the single BPC at the end of the module. Mini-CEX scores changed significantly over time in the 2016 class, but not in the 2017 class. This may be due to changes in departmental assessment practices rather than a true improvement in scores over time. Correlations between the individual WPBA components (average mini-CEX, logbook quota, case presentations) and the final combined paediatric mark were high. The average mini-CEX and case presentation scores were moderately correlated with the single BPC scores. Correlation between the percentage of logbook quota completed and clinical scores were however weak. Due to the low failure rate in this cohort, predictive values for struggling students could not be determined. The continuous formative WPBA practices were found to be both valid and reliable when using Kane's (2013) and Royal's (2017) frameworks for analysis.

The continuous formative WPBA practices in the Department of Paediatric Dentistry at the University of the Western Cape have proven to be feasible as it is integrated in the daily routine patient care provided in the paediatric dental clinics. The findings of this study suggests that the continuous formative WPBA scores are an indication of clinical competence of final year dental students at UWC and could be used to decide if students have reached the expected level of clinical competence in this module. The addition of the single BPC could be reconsidered due to its feasibility and reliability concerns. Further prospective research is however necessary to determine the reliability and validity of the continuous formative WPBA practices.

To our knowledge, this is the first study to evaluate the utility of continuous formative WPBA in making pass/fail decisions in the South African context.

OPSOMMING

Kliniese bevoegdheid is veelsydig en vereis die integrasie van kennis, vaardighede en houdings. Die kliniese omgewing waar pasiënte behandel word, bied 'n geleentheid vir studentevaluering van kliniese vaardighede in 'n egte werkpleksomgewing op die 'doen'-vlak van gedrag.

Finalejaar tandheelkundige studente in die Departement Pediatriese Tandheelkunde aan die Universiteit van Wes-Kaap word daagliks in die kliniese omgewing geassesseer deur middel van verskeie kliniese evaluerings oor die loop van 'n jaar. 'n Bykomende kliniese evaluering aan die einde van die module in die vorm van 'n enkel, blinde pasiëntgeval (BPG) word vereis om te besluit of studente die verwagte vlak van kliniese vaardigheid bereik het.

Beide die betroubaarheid en haalbaarheid van die enkel einde-van-module kliniese geval is bevraagteken in hierdie omgewing. Die nut van deurlopende formatiewe WPBA tydens 'n kursus om progressie te bepaal wat aan die einde van 'n program behaal word, word egter nog in die literatuur bespreek.

Hierdie studie het gepoog om vas te stel of die huidige deurlopende WPBA-uitslae in die Departement Pediatriese Tandheelkunde gebruik kan word as 'n aanduiding van kliniese vaardigheid van finalejaarstudente. 'n Terugwerkende, kwantitatiewe, deursnitstudie is uitgevoer van alle volledige assesseringsrekords van finalejaarstudente (2016-2017) van die Departement Pediatriese Tandheelkundige by UWK. Die kenmerke van, en die verband tussen die deurlopende WPBA komponente, is saam met 'n evaluering van die betroubaarheid en geldigheid van die assesseringsresultate geanaliseer.

Aan die einde van die module het studente oor die algemeen die hoogste telling behaal vir die enkele BPG. Mini-CEX tellings het aansienlik verander in die 2016-klas, maar nie in die 2017-klas nie. Dit kan wees as gevolg van veranderinge in departementeleassesseringspraktyke eerder as 'n ware verbetering in tellings oor tyd. Korrelasies tussen die individuele WPBA komponente (gemiddelde mini-CEX, kwota, gevallestudies) en die finale gekombineerde pediatriese punt was hoog. Die gemiddelde mini-CEX- en gevallestudie tellings was matig gekorreleer met die enkele BPG tellings. Korrelasie tussen die persentasie kwota voltooi en kliniese tellings was egter swak. As gevolg van die lae druipekoers in hierdie kohort, kon voorspellende waardes vir sukkelende studente nie bepaal word nie. Die deurlopende formatiewe WPBA-praktyke

is beide geldig en betroubaar volgens die raamwerke van Kane (2013) en Royal (2017) wat vir analise doeleindes gebruik is.

Die deurlopende formatiewe WPBA-praktyke in die Departement Pediatriese Tandheelkunde aan die Universiteit van Wes-Kaap is haalbaar omdat dit geïntegreer is in die daaglikse roetine pasiëntesorg wat in die pediatriese tandheelkunde klinieke plaasvind. Die bevindings van hierdie studie dui aan dat die deurlopende formatiewe WPBA tellings 'n aanduiding is van die kliniese vaardigheid van die finalejaar tandheelkundige studente by die UWK en kan gebruik word om te besluit of studente die verwagte vlak van kliniese vaardigheid in hierdie module bereik het. Weens uitdagings met uitvoerbaarheid en betroubaarheid, kan die enkele BPG heroorweeg word. Verdere voornemende navorsing is egter nodig om die betroubaarheid en geldigheid van die deurlopende formatiewe WPBA praktyke te bepaal.

Volgens ons kennis, is hierdie die eerste studie in die Suid-Afrikaanse konteks wat die waarde van deurlopende formatiewe WPBA evalueer om slaag/druip besluite te kan neem.

CHAPTER 1: INTRODUCTION

The end result of training health professionals, regardless of the discipline, is to ensure that graduates are clinically competent and can treat patients safely. Teaching, learning and assessment strategies should therefore enable students to develop clinical competence in their chosen discipline, ensuring that competencies have been achieved and can be applied at the end of a course (Hays *et al.*, 2015; Downing & Yudkowsky, 2009; Wass *et al.*, 2001; South African Qualifications Authority, 2001).

Assessment of health professionals measures the outcomes of a course against standards or pre-set criteria as determined by professional regulatory bodies (Wood, 2010; Epstein & Hundert, 2002). The Health Professions Council of South Africa (HPCSA, 2014) has incorporated an adapted version of the 2005 CanMEDS competency framework (The Royal College of Physicians and Surgeons of Canada, 2015) into its guidelines, where competencies are defined as observable and measurable abilities that, when actively integrated in practice, constitute health professional competence (Frank *et al.* 2010).

Assessment of these competencies can be quite complex as it aims to evaluate the integration of knowledge, skills, and attitudes in practice (Downing & Yudkowsky, 2009). The clinical environment allows for authentic patient exposure and learning of professional tasks whilst engaging in patient care. Assessment in the workplace can thus provide an ideal opportunity for assessment of competencies in the context of professional practice (Beard, 2011; Epstein & Hundert, 2002).

Assessment decisions should be sound, defensible and meaningful, and competencies being assessed should be aligned with the intended outcomes (Downing & Yudkowsky, 2009). Our assessment tasks should further be fair and consider the educational impact on our students. Purely summative assessment practices at the end of a programme, where students either pass or fail, may not be beneficial for student growth and development (Schuwirth & van der Vleuten, 2010). Summative assessment can however have formative (for learning) potential if it is used at various points in the programme to determine progression gained up to a certain stage (Wood, 2010). Formative assessment with feedback has been shown to influence student performance positively (Norcini & Burch, 2007), but the value of continuous formative assessment during a course in determining progression gained at the end of a programme is still being debated (Anziani *et al.*, 2008; Riaz *et al.*, 2015).

As with other health professionals, dental students have to demonstrate competence in treating patients by the end of their course. The overall undergraduate dentistry curriculum at the University of the Western Cape comprises six disciplines or modules of which Paediatric Dentistry is one. In order for students to be allowed access to the final examination, which will enable them to graduate from the dental programme, they are expected to pass each of the six modules individually during the fifth/ final year of study. The final integrated examination is in a written, case-based format and includes contributions from all the disciplines within Dentistry; making the need for sound assessment of paediatric clinical competence of final year students critical during the paediatric course.

Continuous workplace based assessment (WPBA) practices have been implemented in the Department of Paediatric Dentistry since 2016. This includes daily direct observation assessments of student-patient interactions in the workplace over the course of a year, the completion of a minimum quota of clinical procedures, and two case presentations. Students are evaluated and receive feedback on all clinical aspects of Paediatric Dentistry including their ability to formulate a diagnosis based on history taking, integration of knowledge, ability to address the patient's main complaint and actual treatment procedures. An end-of-module paediatric clinical assessment is further required to assist with pass/fail decisions in the discipline. This assessment task comprises a single blinded patient case (BPC). The reliability of the single clinical case has been questioned. This assessment task has been challenging to implement due to teaching and assessment demands experienced by the limited staff complement. More than 80 students have to be assessed by three staff members on two learning platforms during the final term of the fifth year undergraduate programme. Due to the disruptions of the recent #FeesMustFall campaign, the Department of Paediatric Dentistry moreover has had to make pass/fail decisions without the end-of-module clinical assessment marks.

There appears to be less literature on the correlation between formative assessments and final grades, with no studies evaluating the utility of continuous formative WPBA in making pass/fail decisions in the South African dental context.

This study aimed to determine whether the current WPBA practices in the Department of Paediatric Dentistry at the University of the Western Cape are an indication of clinical competence of final year students. For this purpose, the characteristics of, and correlation between, the continuous workplace-based assessment (WPBA) components

were analysed together with the reliability and validity of the workplace-based assessment results to ascertain whether inferences drawn from these results could assist the department to make decisions regarding the assessment practices currently employed.

Sound assessment of clinical competencies will be discussed in Chapter 2, followed by a detailed description of the current assessment practices employed in the Department of Paediatric Dentistry at the University of the Western Cape in Chapter 3. The research design and methodology will be clarified in Chapter 4, followed by the results in Chapter 5. In Chapter 6, the key findings will be highlighted and the results discussed in relation to the literature. Implications flowing from the findings will be discussed. Conclusions addressing the specific objectives for the study, with recommendations, will be presented in Chapter 7.

CHAPTER 2: LITERATURE REVIEW

To appreciate the various facets of assessment of competencies of healthcare professionals, this chapter will focus on the purpose of assessment, the principles of sound assessment and the assessment of clinical competence. The review will further emphasize how assessment practices should conform to expected standards. Since the focus of this research is not on the assessment of competencies in higher education in general, this will not be reviewed in detail. Assessment of competencies in the workplace will be the focus of the review.

2.1. Purpose of Assessment

Assessment can be used for baseline, diagnostic, formative or summative purposes (Crisp, 2012). At baseline, assessment can provide information on students' current level of knowledge. It can therefore be diagnostic in nature in that it can be used for feedback or evaluation of student performance and teaching methods, and provide information on changes that need to be implemented for improvement (Crisp, 2012). Assessment can have an "educational impact" (van der Vleuten & Schuwirth, 2005, p.309) (consequential validity) as it can determine how and what students learn. The context in which assessment takes place can therefore also impact on learning (van der Vleuten & Schuwirth, 2005).

Formative assessment or assessment for learning is used to guide the student at various stages of the learning process (Epstein, 2007). It is a frequent, ongoing, active process (Wood, 2010) which highlights where students need to improve by enabling them to recognize the gaps in their knowledge (Black & Wiliam, 1998) and allowing them to assess their level of performance and competence (Schuwirth & van der Vleuten, 2010) with respect to the learning outcomes (Downing & Yudkowsky, 2009).

Feedback is an essential aspect of formative assessment and contributes to the learning process (Riaz *et al.* 2015; Yorke, 2003), as students gain an understanding of where they went wrong and where they still need to improve. The student's performance should be evaluated against specific criteria and feedback should be provided to the student through discussion and allocation of grades (Yorke, 2003). In order to be beneficial, feedback should be explicit, appropriate, meaningful and should be provided timeously (Downing & Yudkowsky, 2009) and in a non-judgemental manner (Wood, 2010). It should also be provided in small chunks (Taylor & Hamdy, 2013; Lara *et al.*, 2016) so

that students are not overwhelmed, and should be relevant, frequent and specific (Black & William, 1998).

Summative assessment informs decisions as to whether a student is fit to progress to the next level or competent enough to qualify (Black & William, 1998; Schuwirth & van der Vleuten, 2010; Wass *et al.*, 2001). Summative assessment usually takes place at the end of a course (Downing & Yudkowsky, 2009; Yorke, 2003) and measures whether the learning outcomes have been met by comparing student performance at the end of a programme against a specified standard. This is therefore an evaluative type of assessment or assessment of learning that has taken place i.e. it “sums up” student achievement (Sadler, 1989, p.120). Summative assessment only provides a snapshot of educational success at a particular point in time and does not necessarily reflect the student’s actual learning experience (Hays *et al.*, 2015). It can however indicate how much learning has taken place (Crisp, 2012).

Schuwirth & van der Vleuten (2010) argued that good assessment practices should make use of varied assessment methods and instruments, using a combination of formative and summative assessment methods. This forms part of an integrated assessment approach which is important when assessing clinical competence (SAQA, 2001).

There is a paucity of literature on the relationship between formative and summative assessment to ascertain the value of formative assessment in determining progression gained at the end of a programme. Azzi *et al.* (2015) showed that formative assessments during a clinical anatomy course, helped to identify students at risk of failing. A moderate correlation between the formative theory assessments and end-of-course mark was found, with a strong correlation between a formative practical assessment and the end-of-course mark. In a retrospective study of undergraduates in oral surgery, Anziani *et al.*, (2008) compared the overall scores for formative and summative assessments. A positive correlation was found between components within the individual formative and summative aspects, but the formative score obtained during the course did not seem to predict the summative end-of-course score.

2.2. Principles of sound assessment

The methods of assessment employed should be an appropriate indicator of student performance and should be designed in such a way so as to measure whether the intended outcomes that are expected of students at the end of a programme have been met (van der Vleuten & Schuwirth, 2005) and are congruent with the competencies being

tested (Downing & Yudkowsky, 2009). Principles that ensure such congruency and relevance include reliability, validity, fairness and feasibility.

Reliability

Reliability refers to the reproducibility of an assessment over a period of time (Downing, 2004). Reproducible and consistent results can be meaningfully interpreted (Downing, 2004). A low reliability of an assessment method implies that measurement errors are prevalent in the assessment data (Downing, 2004).

Reliability goes hand-in-hand with sampling. The sample must be large enough to account for variance across all domains e.g. examiners, patients, content to be tested, instruments in order to minimize errors and improve reliability (van der Vleuten & Schuwirth, 2005). When it comes to assessing clinical competence, the consistency with which the student performs over multiple cases (inter-case reliability) is important (Wass *et al.*, 2001). Multiple observations over time can help to overcome any flaws in the individual assessments (Epstein & Hundert, 2002; van der Vleuten, 1996).

Reliability is usually gauged using a quantitative approach (i.e. inter-rater reliability) where consistency across multiple evaluations and examiners is assessed. Increasing the number of examiners across a range of cases will result in greater inter-rater reliability and consistency between different examiners (Wass *et al.*, 2001). The use of rubrics can further enhance the reliability of assessments through standardisation of criteria and standards (Jonsson & Svingby, 2007).

Rubrics are essentially scoring guides that can either be characterised as holistic or analytic (Perlman, 2003). Holistic rubrics assess the overall student performance using a single scale (Perlman, 2003) where individual components/ categories are not judged separately (Nitko, 2004 as cited by Gezie *et al.*, 2012). On the other hand, analytic rubrics have two or more separate scales or categories (Perlman, 2003) which provide specific feedback for each criterion on the scale (Mertler, 2001; Nicholson *et al.*, 2009). Nicholson *et al.* (2009) found that when compared with analytic rubrics, the judgements obtained with holistic rubrics were more consistent, but analytic rubrics provided more detail regarding where intervention might be needed.

In cases where clinical supervisors are reluctant to provide honest feedback (Norcini & Burch, 2007), or where they feel pressured by the student to give a good score (Beard, 2011), use of explicit criteria makes it easier to provide feedback on observed

performance (Norcini & Burch, 2007). Rubrics also ensure that the assessment is less subjective i.e. consistency of judgement is improved (Jonsson & Svingby, 2007). The issue of subjectivity can be addressed by ensuring that the rubric is well-constructed and that examiners are adequately trained and understand how to use it (Perlman, 2003). They should strictly adhere to the criteria to ensure subjectivity does not influence the assessment (Perlman, 2003). It is however important to note that some degree of variability between clinical supervisors will always exist despite extensive attempts at calibration (Govaerts & van der Vleuten, 2013). Interpretation of a rubric adds a qualitative dimension to the assessment and can therefore pose a challenge where reliability of an assessment is concerned (Driessen *et al.*, 2005).

Driessen *et al.*, (2005) suggested an alternative qualitative approach to evaluating the reliability of assessment tasks which are based on supervisor judgement by including markers for credibility and dependability. Credibility or trustworthiness of an assessment has to be supported by evidence and this can be achieved through three strategies, namely, triangulation (i.e. combining information from different sources), prolonged engagement over time and member checking which includes student discussion and feedback. Dependability refers to the quality assurance processes that are put in place (Driessen *et al.*, 2005).

Validity

Validity is a measure of whether or not an assessment task or instrument measures what it is supposed to measure i.e. is it authentic? (van der Vleuten & Schuwirth, 2005). It is important to have clarity on what is being assessed (i.e. knowledge, skills, content, behaviour) (SAQA, 2005) and how it will be assessed (method of assessment) (Epstein, 2007; Wass *et al.*, 2001).

The assessment technique chosen should therefore be specific and appropriate to ensure that it actually measures the desired competencies (Schuwirth & van der Vleuten, 2010). As it is not possible for a single assessment technique to measure all aspects of competence, validity can be improved by making sure outcomes being assessed are explicit and that assessment methods are varied and fit for purpose (van der Vleuten & Schuwirth, 2005).

It could be argued that in order for students in the health professions to be considered competent, they should have the knowledge needed to be able to solve complex presenting problems. Valid assessment, therefore, should provide the student with the

opportunity to demonstrate skills in clinical reasoning which will involve the integration of concepts (Epstein & Hundert, 2002; van der Vleuten & Schuwirth, 2005). Workplace-based assessment (WPBA) takes place in the actual workplace environment. It has therefore been shown to be a good indicator of clinical competence (Beard, 2011) as it requires integration of knowledge and skills (Downing & Yudkowsky, 2009).

Validity also refers to how the assessment is interpreted and if meaningful inferences can be made regarding the appropriateness and usefulness of an assessment (SAQA, 2005; Cook *et al.*, 2015; Kane, 2013). In a workplace-based setting, the interpretations that are drawn should be informed by the context in which the assessment takes place (Govaerts & van der Vleuten, 2013). Inferences made from assessment results, such as interpretation of scores, need to be defended with evidence that will support the decision that is taken on whether or not a student is competent to progress to the next level (Cook *et al.*, 2015; Kane, 2013; Royal, 2017). Multiple sources of evidence are required to prove whether claims of validity of an assessment task can be supported (Downing, 2003).

Much of the literature on validity focuses on written assessments, multiple choice questions and the generation of test scores which are often used to assess theoretical knowledge. Content validity is applicable in these cases as it is important to ensure that tests and examinations cover the entire course content (Schuwirth & van der Vleuten, 2010) and aligns with the learning outcomes (Wass *et al.*, 2001). Blueprinting involves the mapping of course content and planning of assessment activities to ensure that all the learning outcomes are sampled (Patil *et al.*, 2015). It is also important that an assessment instrument is able to discriminate between students who function on a higher cognitive level and those who do not. This is embodied in the concept of construct validity (Schuwirth & van der Vleuten, 2010). Downing (2003) posited that “all validity is construct validity” (p.831) as it encompasses all facets of validity.

Due to the variable nature of WPBA, content validity as mentioned above, is difficult to apply to the workplace-based setting where a range of patients with varied needs are treated. To overcome this challenge, studies have compared the scores of various clinical assessment tasks and methods to support final assessment decisions (Durning *et al.*, 2002; Hatala *et al.*, 2006; Boulet *et al.*, 2002). In all three studies (Durning *et al.*, 2002; Hatala *et al.*, 2006; Boulet *et al.*, 2002), the mini-CEX was compared with other assessments like specialty examinations (Hatala *et al.*, 2006), standardized patient (Boulet *et al.*, 2002) and monthly evaluation forms (Durning *et al.*, 2002). Strong

correlations between these different assessment scores have been used as a measure of validity of decisions (Durning *et al.*, 2002; Boulet *et al.*, 2002).

The reliability and validity of scores generated from clinical assessments have however been questioned due to the perceived subjective professional judgement of the assessors (Govaerts & van der Vleuten, 2013). This interpretation could however be seen as a valuable source of qualitative information that can help to inform decisions on the interpretation of scores rather than a potential for an increase in measurement error (Govaerts & van der Vleuten, 2013). Govaerts & van der Vleuten (2013) argue that all perceptions should be considered valid and can add value to the overall observations regarding clinical competence.

A move away from statistical analysis of individual assessment methods or tasks to determine the validity of decisions has become part of the recent discourse on assessment (Harris *et al.*, 2017). Modern validity theories rather focus on the meaningful interpretation of data. Various frameworks have been proposed on how to structure the necessary evidence for making a case for validity of an assessment (Downing, 2003; Kane, 2013; Royal, 2017). The sources of evidence suggested by Downing (2003) are more suited to written tests and examinations, but, frameworks put forward by Royal (2017) and Kane (2013) can be applied to clinical contexts and are described below.

Royal (2017) proposed four tenets for the evaluation of validity. These include:

1. Validity refers to inferences, not instruments i.e. scores generated by instruments should be interpreted critically.
2. Validity evidence, interpretation and use: More than one source of evidence should be used to support claims of validity and it should be interpreted within the appropriate context.
3. Validity is a continuum i.e. validity can vary between assessment components.
4. Validation is an ongoing process (as the learning/ assessment environment is dynamic).

Kane (2013) identified four inferences. These include:

1. *Scoring*: This refers to how clinical supervisors translate their observations into a score. There is some overlap between this point and the first tenet proposed by Royal (2017).

2. *Generalisation* deals with the test and factors that influence it. Multiple sources are considered including content that is covered and number of raters/ observations/ cases.
3. *Extrapolation* of the test score to the real world environment. For assessments that take place in the workplace-based environment, extrapolation of results to the real-world setting does not apply.
4. *Implications* of interpreting the inferences that can be drawn from these interpretations can inform a decision regarding a student's clinical competence.

As will be presented in Chapters 5 and 6, both Royal's four tenets (2017) as well as Kane's four inferences (2013) could be applied to the case for the validity of WPBA inferences made in this study.

Fairness

According to SAQA (2005), fairness in assessment ensures that the learner will not be disadvantaged in any way. In order to promote fairness, the assessment process should be transparent and objective and should be based on known outcomes (SAQA, 2005). The demands made should be achievable and explicit, and clear standards/ criteria should be set in order to measure the actual level of knowledge of each individual student (Downing & Yudkowsky, 2009). Making scoring rubrics available to students can clarify expectations and contribute to fairness (Jonsson & Svingby, 2007; Coderre *et al.*, 2009).

Each individual examiner brings their own conceptions and perceptions to the assessment space (Tziner *et al.*, 2005). Factors that can negatively predispose an examiner towards a student could be considered bias and may be based on gender, ethnicity, language or the examiner's innate personality or attitude (McManus *et al.*, 2013; Stupart *et al.*, 2008). This problem can be avoided by having more than one examiner assessing the same student (McManus *et al.*, 2013). However, in a resource-constrained setting with limited staff, this might not be possible.

The concept of fairness is also transferred to the learning environment as nervousness or test conditions could influence student performance and negatively affect the reliability of an assessment, especially in high stakes assessments (van der Vleuten & Schuwirth, 2005; Wass *et al.*, 2001). One of the advantages of continuous assessment is that it is associated with reduced levels of anxiety when compared with high stakes formal examinations at the end of a programme (Sadler, 1989).

Feasibility

The feasibility of an assessment refers to whether it can be implemented successfully. Factors such as staffing required to implement a particular assessment activity, other resources such as time, infrastructure, equipment and cost, can influence whether the assessment task/ activity is realistic and achievable (Schuwirth & van der Vleuten, 2010). In other words, these factors determine whether the assessment is feasible. Some activities can be labour and resource intensive (Wass *et al.*, 2001). These challenges, together with the setting and specific circumstances surrounding the assessment activity should be taken into consideration before an assessment activity is chosen.

In summary, ensuring sound assessment practices is a complex process. Assessment tasks should test different levels of comprehension and take any possible causes of bias on the part of the examiners into consideration (Patil *et al.*, 2015). The final assessment score obtained should be representative of students' abilities, regardless of the number of assessment tasks they may be subjected to.

Even though an assessment programme may consist of a range of varied assessment tasks, emphasis is increasingly being placed on holistic evaluation of all assessment components in a particular programme i.e. programmatic assessment (van der Vleuten *et al.*, 2015). Programmatic assessment should be based on an overarching structure or competency framework which stipulates the details of the assessment activities and how a decision on competence of the student is reached (van der Vleuten *et al.*, 2015). Feedback is an essential aspect of programmatic assessment (van der Vleuten *et al.*, 2015). All tasks have to be evaluated as a whole in order to make a decision on whether a student is fit to pass. The whole is therefore more valuable than the individual components as the credibility of the final decision is based on information obtained from a variety of sources (van der Vleuten *et al.*, 2015).

By making sure that assessment practices are comprehensive and explicit, and by improving the precision of assessment practices, more reliable conclusions can be drawn regarding student competence and promotion.

2.3. Assessment of Clinical Competence

Clinical competence encompasses knowledge, skills, professionalism (Hays *et al.*, 2015; Epstein & Hundert, 2002) and in particular, diagnostic problem-solving skills which involve data gathering and diagnosis.

A key learning theory relating to assessment was proposed by Miller (1990), and is depicted as a pyramid (Figure 2.3.1).

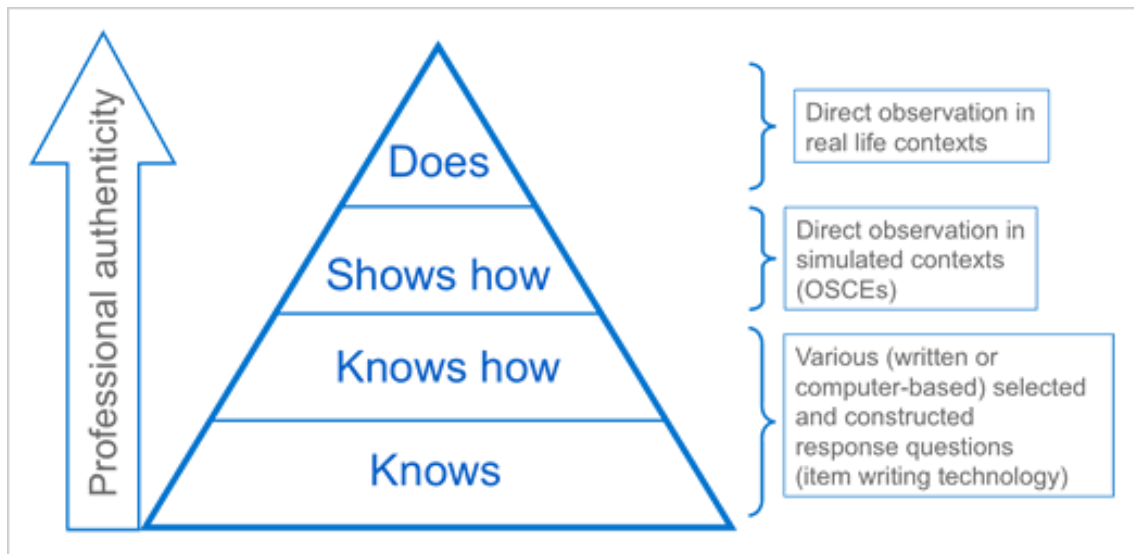


Figure 2.3.1: Adaptation of Miller's Pyramid (Miller, 1990; Wass, 2001)

All levels of Miller's pyramid play some role in developing clinical competence (Wimmers, 2006). In order to be considered clinically competent, a student or health professional must not only have the knowledge (which forms the base of Miller's pyramid) but they should also be able to apply that knowledge to clinical situations (tiers 2 and 3) so as to be able to manage or treat a patient effectively (top tier). The uppermost tier of Miller's pyramid is synonymous with integration and higher-order thinking (Downing & Yudkowsky, 2009; Miller, 1990). The ability to solve a problem is however, case-, content- and context-specific (van der Vleuten, 1996; Epstein & Hundert, 2002; Adamson *et al.*, 2012; Epstein, 2007). Similarly, whether or not a student can be considered clinically competent is also case-dependent (Wimmers, 2006). Students should, therefore, be able to apply the theoretical concepts to the various clinical scenarios they will encounter during their professional careers (Morrison & Free, 2001; Epstein, 2007).

Clinical competence can be evident through direct observation or video review (Epstein, 2007) of skills like history taking, patient examination and patient management (Wimmers, 2006). Observation is invaluable when assessing basic clinical skills (Norcini & Burch, 2007). According to Tanner (2006), clinical judgement encapsulates four domains i.e. "noticing, interpreting, responding and reflecting" (p. 208), all of which are essential when formulating a diagnosis.

The concept of workplace-based assessment was introduced in Chapter 1. Workplace-based assessments are carried out in the workplace environment and are valuable in assessing actual clinical performance at the 'does' level of Miller's pyramid (Beard, 2011; Harris *et al.*, 2017). WPBA practices are characterised by feedback, and include reflection and skills development to encourage lifelong learning (Harris *et al.*, 2017).

Observation of students in the workplace provides an ideal opportunity to assess clinical competence and provide feedback, thereby facilitating learning (Beard, 2011) and allowing students to improve on their performance (Norcini & Burch, 2007; Epstein, 2007). This formative type of assessment with timely and specific feedback has been shown to result in positive behaviour change among students (Norcini & Burch, 2007).

WPBA is not without its challenges though. Studies have shown that it is not used as often as it should be, and that feedback and reflection are often lacking (Norcini & Burch, 2007). It is further regarded by some with cynicism due to the unpredictable and unstandardized assessment tasks, subjectivity of assessors and biased performance ratings (Albanese, 2000; Downing, 2005; Kreiter & Ferguson, 2001). Although these pose a threat to the reliability and validity of WPBA, reliability is improved by multiple assessments (Norcini *et al.*, 2003). These assessments are also usually integrated, thereby improving validity (van der Vleuten & Schuwirth, 2005) as has been described previously.

Workplace-based assessment methods include Clinical Encounter Cards (CEC), Blinded Patient Encounters (BPE), Direct Observation of Procedural Skills (DOPS), Case-based Discussions (CbD), and the Mini-Clinical Evaluation Exercise (mini-CEX), (Norcini & Burch, 2007). All these assessment methods have a feedback component based on specified criteria and are considered formative in nature (Norcini & Burch, 2007).

Clinical encounter cards (CEC) involve direct observation of clinical skills using a six-point rating scale. Information regarding the quality of the performance as well as feedback provided to the student are recorded on score cards (Norcini & Burch, 2007). Hatala and Norman (1999) showed this to be a feasible, valid and reliable tool provided that at least eight encounters are considered to ensure reliability. The detailed feedback process with this tool led to greater student satisfaction (Paukert *et al.*, 2002).

Blinded patient encounters (BPE) are based on the same principles as the CEC and mini-CEX. In this situation, the patient is unknown to the student who is assessed on

problem-solving skills and the formulation of a diagnosis. As with the mini-CEX encounter, a nine-point rating scale is used to generate feedback (Norcini & Burch, 2007).

Direct observation of procedural skills (DOPS) allows for assessment of actual clinical procedures. As is the case with logbooks, DOPS also records the procedures completed. It is thus similar to a logbook which is designed to ensure that the minimal procedural requirements have been met, however, the latter is usually not associated with feedback (Norcini & Burch, 2007).

With case-based discussions (CbD), the student presents the records of a selected patient case to an assessor for discussion purposes. Students are assessed on clinical reasoning and the rationale behind the decisions which are made (Norcini & Burch, 2007).

The mini-CEX has been widely used in the workplace-based environment (Pelgrim *et al.*, 2011) and can be applied to a variety of settings (Norcini & Burch, 2007). It is an abbreviated version of the traditional clinical examination (CEX) and can be completed within 15 to 20 minutes (Norcini *et al.*, 2003). These mini-CEX sessions are single patient encounters (Pelgrim *et al.*, 2011) which are conducted over a period of time (Norcini *et al.*, 2003).

Being able to conduct a thorough clinical examination and formulate a diagnosis is one of the foundations of being a good clinician (Norcini & Burch, 2007). The mini-CEX has been used to assess clinical competence through observation of the student-patient interaction and can include history taking, formulation of a management plan or carrying out of specific clinical tasks (Al Ansari *et al.*, 2013; Norcini *et al.*, 2003; Norcini & Burch, 2007). The staff member then assesses the student and provides feedback on performance so as to guide the student to meet the desired learning outcomes at the end of the programme (Norcini & Burch, 2007).

Structured feedback (using rating scales) is a characteristic of the mini-CEX encounter. Feedback is provided throughout the duration of student training for different clinical situations (Norcini & Burch, 2007). This is especially important where workplace-based assessments are concerned (Beard, 2011) and where more than one examiner is involved in the assessment (Adamson *et al.*, 2012). Norcini & Burch (2007) highlighted the value of clearly defined criteria to facilitate objective feedback based on observations.

The mini-CEX makes use of criteria which have been proven to be valid (Al Ansari *et al.*, 2013). However, even though Pelgrim *et al.*, (2011) agreed that the mini-CEX was one of the best-supported instruments, they emphasized that more evidence of construct validity was needed. Multiple mini-CEX assessments conducted over a period of time improves reliability and validity as student performance across a range of patients can be observed by different examiners (Al Ansari *et al.*, 2013; Norcini *et al.*, 2003). Interactions with different patients and assessors provide multiple opportunities for feedback (Norcini *et al.*, 2003) and is therefore a valuable formative assessment exercise. Different supervisors will be able to observe a range of skills over a period of time, thereby gaining a reasonable idea of the student's abilities (Norcini *et al.*, 2003). This is essential when evaluating the level of clinical competence (Wass *et al.*, 2001) as inferences made from multiple observations over time give a more accurate picture of competence to the point where a summative decision can perhaps be supported (Harris *et al.*, 2017).

The WPBA tools discussed above (i.e. mini-CEX, BPE, DOPS and Cbd) all have bearing on the Paediatric Dentistry assessment and will be discussed within this context in the subsequent chapter. Currently, there is a paucity of literature regarding the application of these workplace-based assessment tools in the dentistry setting nor have conclusions been drawn from assessment results utilizing these tools. The value of these assessments in making pass/fail decisions in this particular context has also not been explored in the literature. The present study will attempt to add to this body of knowledge.

CHAPTER 3: CURRENT ASSESSMENT PRACTICES FOR THE FINAL YEAR PAEDIATRIC DENTISTRY MODULE

This chapter describes the current assessment practices for the final year Paediatric Dentistry module at the University of the Western Cape; focusing on the clinical component of the assessment plan. The educational theory that informed the current clinical assessment tasks will be highlighted.

3.1. Assessment Plan

As mentioned previously, Paediatric Dentistry is one of the six modules in the dentistry curriculum that undergraduate students have to pass in the fifth/ final year of study. Students are expected to achieve 50% in each of the six modules individually in order to be allowed access to the final examination, which will enable them to graduate from the dentistry programme. This final integrated examination covers content from all the disciplines within Dentistry.

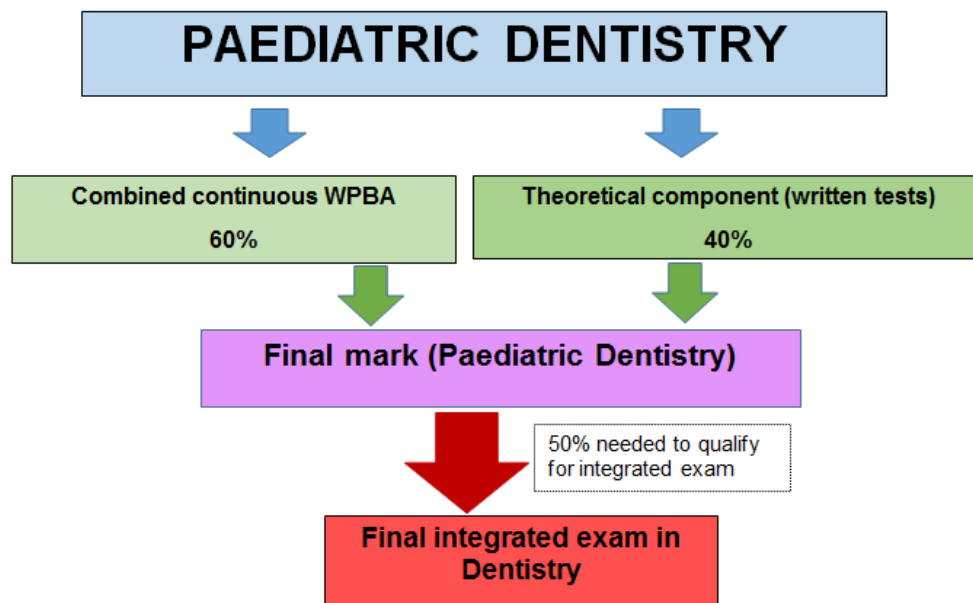


Figure 3.1.1: Overall assessment plan (Paediatric Dentistry)

Assessment in the discipline of Paediatric Dentistry takes place in the clinical setting as well as in tutorials and traditional test and examination environments. In the final year, greater emphasis is placed on continuous assessment which includes a clinical component comprising of various workplace-based assessment (WPBA) tasks, as well as written tests on theoretical content. As a series of tasks are evaluated, it enables a

broader assessment of skills and knowledge. The clinical component carries a higher overall weighting of 60% compared to 40% for the theoretical component (Figure 3.1.1). As the focus of this research is on clinical competence, the theoretical component will not be discussed further.

The clinical component itself is divided into various WPBA tasks; all of which contribute to a varying degree to, what is referred to as the 'combined continuous workplace-based assessment' score (Figures 3.1.2 and 3.1.3). WPBA's take place on a daily basis in the undergraduate Paediatric Dentistry clinics. Clinical and theoretical knowledge is assessed through observation of patient management, together with motivation for why students choose a particular treatment option. Assessment is therefore integrated and engages the top tiers of Miller's pyramid (Miller, 1990). WPBA tasks include direct observation assessments of student-patient interactions, the completion of a minimum clinical quota of procedures, and two case presentations.

These continuous WPBA tasks are formative in nature, although marks are given for each assessment task; which contributes to the final paediatric mark. Where formative assessment is typically not associated with the allocation of a grade (Sadler, 1989), continuous assessment can have a formative and summative function where students are graded, but with feedback provided at the same time (Hernández, 2012).

For the purposes of this study, the direct observation assessments of student-patient interactions will be referred to as mini-CEX encounters even though a rubric relevant to the Paediatric Dentistry context is used in place of the original mini-CEX scoring sheet published by Norcini & Burch (2007). Similar to the implementation of the mini-CEX in other clinical settings (Al Ansari *et al.*, 2013; Norcini *et al.*, 2003; Norcini & Burch, 2007), a student's performance is assessed by various staff members through direct observation whilst treating a range of patients. As most of the module credits are devoted to clinical time in the clinical workplace-based setting, the average of the mini-CEX encounters is calculated at the end of the year and contributes 50% to the combined continuous WPBA.

The mini-CEX episodes are directly linked to procedures which form part of an expected clinical quota (Appendix A). The achieved procedures are recorded in a quota logbook and contributes 20% to the combined continuous WPBA score. Two case presentations contribute 10% to the combined continuous WPBA score.

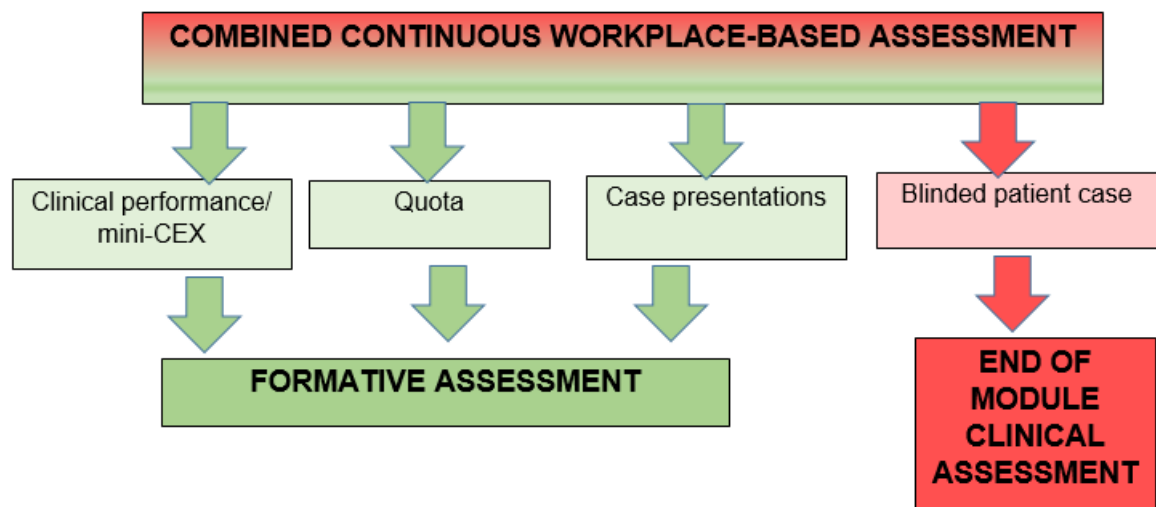


Figure 3.1.2: Clinical assessment components (Paediatric Dentistry)

In addition, a single blinded patient case (BPC) is used to gauge student progress at the end of the module and to assess whether the learning outcomes have been met. This assessment does not include feedback to students. The blinded patient case also contributes to the combined continuous WPBA score; and carries a 20% weighting.

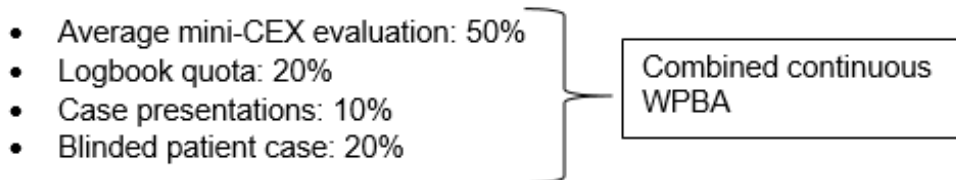


Figure 3.1.3: Weighting of assessment components

Students are required to achieve 50% for the overall paediatric clinical assessment i.e. combined continuous WPBA score (Figure 3.1.1), as well as obtain a combined 50% for the mini-CEX and logbook quota. In addition, they are not allowed to obtain less than 50% in more than two of the individual WPBA components. This would constitute a fail and the student would not be allowed to write the final integrated dentistry examination. They would then have to repeat the final year.

This assessment approach is delineated in the module descriptor (Appendix B) and study guide that students have access to. The implementation of the assessment tasks are described in detail in the following section.

3.2 Implementation of continuous assessment tasks

3.2.1. Mini-CEX evaluations and logbook quota

The department accommodates approximately 80 to 85 final year students in the clinics every year over a period which spans roughly 30 weeks. Each student sees an average of 10 patients for the year during this time where all the treatment the patient requires, is completed i.e. from diagnosis to completing of all treatment procedures.

Routine dental treatment takes place during the four hours of clinical time allocated to each student per week. This includes all diagnostic, restorative, prosthetic, orthodontic and surgical procedures for the child patient. At least half of the patients seen are first-time patients requiring a diagnosis and treatment plan. Students treat a minimum of two patients per week either under local anaesthesia, general anaesthesia or sedation where comprehensive dental treatment is provided under supervision. This provides the ideal opportunity to assess students performing varied professional tasks in a range of workplace-based environments.

Due to the fact that the department is short-staffed, nine part-time staff members are employed to assist with undergraduate clinical teaching and assessment in the clinics. Permanent staff are responsible for all other assessments. All staff members are rotated through the clinics and students are thus exposed to a number of different staff members during the year. Each staff member supervises up to 8 students at a time. Each of these students receive feedback at the end of the 2-hour clinical session. Even though the clinics are extremely busy, staff still manage to record comments and provide feedback.

With the daily mini-CEX evaluations, as previously explained, a holistic clinical rubric with specific criteria (Appendix C) is used to assess professionalism, clinical ability, knowledge and application of knowledge as well as diagnostic and problem-solving skills. This is similar to scoring categories used in clinical encounter cards and the mini-CEX tools as previously discussed in Chapter 2 (Norcini & Burch, 2007). All staff were trained to use the rubrics for clinical assessment, prior to its implementation at the start of the 2nd semester in 2016.

As argued by Perlman (2003), a holistic rubric such as this one aims to assist with standardisation between staff members on how students are assessed in the clinics. It is used to assess everyday patient management from diagnosis and treatment planning to the completion of actual clinical procedures. The student is not assessed on individual categories but rather on overall performance i.e. holistically. The rubric criteria also pay

attention to detail. Based on course requirements, a detailed module descriptor was drawn up (Appendix D) and used as a basis for compiling the rubric which includes evaluating each step of an actual clinical procedure such as placement of a restoration. Incorporating detailed, clearly articulated steps in sequential order for a particular competency can guide the learning process as it gives students an idea of what the expected standards of practice are (Harris *et al.*, 2017). Each category stipulated on the rubric has a number of detailed criteria which are easy to identify in the clinical setting. Any of these criteria could be present in a particular case. Professionalism is only evident in the lowest two categories and distinction is made between totally unacceptable behaviour and a smaller transgression in category two. All categories above the second category assume an acceptable level of professionalism (Appendix C).

Students are assessed on their overall performance and given a score between 1 (lowest) and 5 (highest) for both a clinical and a theoretical component. Prior to the implementation of the rubric, only a clinical mark was recorded and this was based on the supervisor's overall opinion of the student's ability during the session and not on detailed criteria. Scores merely ranged from "unacceptable" (1) to "excellent" (5). This resulted in students slipping through the cracks even if their theoretical knowledge was not up to par. The rubric currently being used therefore gives a more holistic idea of student performance as two marks are recorded on the rubric- one for clinical performance and one for knowledge and insight (theory and application). The average score (at the end of the year) is recorded and rounded up to the next whole number which is then converted into a percentage according to predetermined weighting scales as decided on by the department and as based on the HPCSA competencies for Dentistry. Converting the rubric scores into grades and determining the appropriate scoring levels poses a huge challenge when constructing a rubric and should be based on logic rather than a mathematical formula (Gezie *et al.*, 2012). Likewise, the weighting used to determine the categories for the rubric was not based on a formula. Each score between 1 and 5 represents a percentage from 20% to 75+% and corresponds with the expected competencies as delineated in the rubric.

Even though this rubric is quite detailed, staff reported it to be less time-consuming as scoring is fast which is essential in a busy workplace-based environment. Moreover, it makes it easier to capture the daily clinical marks as opposed to calculating the average percentage for different categories for each patient that a student treats as would be the case with an analytic rubric.

The implementation of the clinical rubric in its current format has been shown to be feasible based on both staff and student feedback. The criteria for evaluation are explicit and staff are required to provide feedback (in keeping with the principles of formative assessment) after each clinical session. Staff members discuss each student's performance with them. However, in order to be beneficial, students should understand why certain standards were not met and how to improve on future attempts (Hays *et al.*, 2015; Wass *et al.*, 2001; Yorke, 2003). Specific verbal feedback relevant to the clinical session is provided. Students are required to reflect on their performance. They are asked to highlight what they did well and where they feel they can improve. This approach is in line with recommendations from the literature where emphasis is placed on development of an "action plan" which provides specific information on how students need to proceed to meet the learning outcomes (Norcini & Burch, 2007, p. 866).

Staff in the Department of Paediatric Dentistry are also required to record comments on the rubric to provide more detail regarding student performance. This is in agreement with Harris *et al.*, (2017) and Govaerts *et al.*, (2013) who encouraged this addition of a narrative as opposed to merely recording a student's mark. These narratives allow scores to be interpreted more meaningfully (Govaerts *et al.*, 2013). The grade allocated by the clinical supervisor is discussed with the student and both parties are required to sign the rubric as acknowledgement that this discussion has taken place. This improves accountability of the supervisor as well as the student and it is the department's way of ensuring "sustainable assessment and feedback strategies" as recommended in the literature (Norcini & Burch, 2007, p.869). Students are aware that the marks allocated for each clinical session count towards their final clinical mark. Anecdotal evidence in the clinics suggests that this knowledge alone encourages them to prepare better for their next clinical session. Similar findings were reported by Riaz *et al.*, (2015) where formative assessment resulted in a positive attitude towards learning.

Clinical practice has to be evaluated in conjunction with logbook experience (Beard, 2011) in order to calculate the final clinical mark. Each procedure has a specific treatment code which is recorded on the rubric, thus making it possible to correlate the codes to the logbook quotas completed. Students receive feedback for every procedure recorded in the logbook. A minimum procedural quota is used as a benchmark for students to aim towards (Appendix A). However, students are often not able to fulfil all the specific requirement due to various external factors that are outside departmental control. These include poor patient compliance that influences the type of treatment that can be provided, poor patient attendance, treatment costs in favour of more cost-effective

treatment options or parental preference for a particular procedure such as requesting extractions over restorative procedures. Patil *et al.*, (2015) emphasised the importance of weighting course content to determine essential knowledge and what is good to know. Similarly, quotas are also weighted by the department according to importance as guided by the HPCSA competencies for dentistry. Procedures involving primary teeth carrying the highest weighting. Rare or less common procedures are allocated a lower weighting.

If a particular quota cannot be obtained, students can be credited with an additional procedure in a different category. The #FeesMustFall campaign has forced management to question the value of logbook quotas and how it correlates with actual clinical competence. Currently, some departments feel students are not competent unless they have met every single quota on the list of requirements even though realistically, it is not achievable. Increasingly, pressure is being placed on departments in the Dentistry faculty to relook at this assessment task as part of the assessment of clinical competence.

3.2.2. Case presentations

Students are required to present their patients seen in the workplace as part of a case for discussion. They motivate their treatment choices and receive feedback from their peers and supervisors. This is seen as a continuation of workplace-based assessment as real patient cases are discussed (Norcini & Burch, 2007). The assessment focusses on clinical reasoning and the rationale behind the clinical choices (Norcini & Burch, 2007), thus honing critical thinking skills (Popil, 2011). Through presentations of their peers' patient cases, students are also exposed to a range of real-world patient cases that they would normally not have had access to (Popil, 2011).

3.2.3. Blinded Patient Case (BPC)

As with the Blinded Patient Encounters (BPE) discussed earlier in Chapter 2 (Section 2.3), students are assessed using a blinded patient case (BPC) that is unknown to them. This assessment happens at the end of the module where their ability to formulate a diagnosis based on history taking, integration of knowledge and ability to address the patient's main complaint is evaluated.

The once-off BPC has multiple challenges in our setting. Students are assessed by one examiner using an analytic rubric (Appendix E). This rubric outlines specific criteria which only focus on diagnosis and treatment planning skills i.e. history taking, diagnosis and formulation of a treatment plan, rationale for the treatment plan and knowledge and insight. Each of these four categories is evaluated separately and scored out of 10. An

average of the components is used to calculate the final mark which is expressed as a percentage. It is up to the discretion of the examiner to allocate a mark below 4 or above 8 based on exceptional clinical circumstances for a particular case. Calculation of the score for the blinded patient case is reported by staff to be more cumbersome than the mini-CEX evaluations which involves encapsulating a holistic impression of student performance in a single score.

Examiners are further removed from the service clinics for the BPC assessment which increases the supervision workload on the rest of the staff who already have to cope with large student numbers. The patients used for the blinded patient case are not standardised, are unpredictable and have to be rescheduled if booked patients fail to attend their scheduled appointments. In addition, as children younger than 12 years of age are used in these assessments, expecting them to sit for more than one student examination is not feasible as they tire easily and have a limited attention span. Different patients are therefore booked, thereby threatening the fairness of this assessment.

At present, there are only three permanent staff members who are responsible for this assessment and they have to cater to approximately eighty final year students on two platforms i.e. Mitchell's Plain and Tygerberg campuses over a short period i.e. the final term. Questions have therefore been raised regarding the feasibility and reliability of this assessment task.

The Department of Paediatric Dentistry uses a variety of assessment methods to determine whether a student can be considered competent and be allowed to write the final integrated examination in Dentistry. Feedback is provided to enable student growth and development (van der Vleuten *et al.*, 2015). All tasks are mapped according to an overarching framework (Figures 3.1.2 & 3.1.3) which clearly stipulates the contribution of each assessment component to the final combined continuous WPBA score. Challenges regarding feasibility, reliability and validity of assessment practices however remain; with this study aiming to determine if the current WPBA practices are an indication of clinical competence of final year dental students in the Department of Paediatric Dentistry at the University of the Western Cape.

CHAPTER 4: RESEARCH DESIGN AND METHODOLOGY

4.1. Problem statement

Continuous formative WPBA practices have been implemented in the Department of Paediatric Dentistry at the University of the Western Cape since 2016. An end-of-module clinical assessment is an additional requirement to inform pass/fail decisions. The reliability and feasibility of the single clinical case (BPC) at the end of the module have however been questioned. To our knowledge, continuous WPBA practices as an indication of clinical competence have not been described in the South African dental setting before. This study will evaluate the utility of the continuous formative workplace based assessment scores in assessing clinical competence in final year dental students in the Department of Paediatric Dentistry at the University of the Western Cape. These findings could inform sound assessment practices.

4.2. Research Question

Are continuous WPBA practices an indication of clinical competence of final year students in the Department of Paediatric Dentistry at the University of the Western Cape?

4.3. Aim and Objectives

The overall aim of the study was to determine if the current continuous WPBA practices in the Department of Paediatric Dentistry at the University of the Western Cape are an indication of clinical competence of final year students.

Objectives:

1. To determine the correlation of scores between the various components of the continuous WPBA.
2. To determine the predictive value of the continuous WPBA scores in identifying struggling students.
3. To determine the reliability of the WPBA results.
4. To determine the validity of the WPBA results.

4.4. Study design

A retrospective, cross-sectional study was conducted using both quantitative and qualitative approaches.

4.5. Setting

The study was conducted in the Department of Paediatric Dentistry at the University of the Western Cape. The current staff:student ratio is 1:8. In a discipline which involves actively providing all types of dental treatment to children under the age of 12 years, the ratio should ideally be 1:5 or 1:6, especially considering that most treatment takes place under local anaesthetic and that behaviour management can be quite a labour-intensive and time-consuming part of clinical supervision and patient management. A major part of the clinical sessions is the service delivery component where approximately 160 children are treated each week. As there are only three full-time staff members in the department, nine part-time general dental practitioners are employed to assist with the clinical teaching load on two clinical platforms (Mitchell's Plain and Tygerberg) which cater to roughly 160 students over the 4th and 5th years of study. Approximately 80 of these are final year dental students.

4.6. Target and Study Population

The target population consisted of final year students from the Paediatric Dentistry Department at the University of the Western Cape over a two-year period (2016- 2017). Study participants were identified from routinely kept university records.

Inclusion criteria:

- i. Complete set of scores for all the continuous workplace-based assessment components.
- ii. All final year students were included, irrespective of final fail/ pass/ repeat assessment results.

Exclusion criteria:

- i. Incomplete records

4.7. Data collection and analysis

Data sources

Clinical assessment scores (for all assessment opportunities) were retrieved from departmental records. The theoretical component of the assessment was not included.

Data sources used were:

1. Completed mini-CEX rubrics (hard copies)
2. Routinely compiled annual departmental spreadsheets containing all clinical components of students' assessment marks as explained in Figures 3.1.2 and 3.1.3.

Data collection and management

All data (including rubrics and spreadsheets) were handled and managed according to Good Clinical Practice (GCP) requirements and ethical standards. Students whose assessment scores were included in the study were assigned a unique student identifier number. This data was entered into an electronic database (an Excel spreadsheet). Student names and student numbers were not entered into the electronic database.

The score sheets were stored in a locked filing cabinet and all electronic data was stored on a password protected computer to which only the researcher had access. All paper documents will be kept in a locked cupboard for a period of 5 years.

The electronic databank was backed up on a daily, weekly, monthly and 6-monthly basis onto a secured storage disk. Copies were used to perform calculations and analysis.

Variables collected for individual students:

- Both the individual and average mini-CEX scores.
- Blinded patient case scores
- Percentage of logbook quota completed
- Scores for clinical case presentations
- Combined continuous WPBA scores

Operational definitions (please also refer to Chapter 3)

For the purpose of this study, the following definitions were used:

- Mini-clinical evaluations (mini-CEX)- Daily direct observation assessments of student-patient interactions adapted from the mini-CEX as described by Norcini & Burch (2007). The average of this mark for all encounters was calculated at the end of the year to obtain the average mini-clinical evaluation score (%).
- Logbook quota completed- The percentage of the recommended quota for the various clinical procedures that has been attained.
- Clinical case presentations- discussion of patient assessment and treatment choices
- Blinded patient case- The end-of-module clinical examination, assessing the student's ability to formulate a diagnosis based on history taking and assessment of a patient.
- Combined continuous WPBA score- Overall end-of-module score obtained. This mark includes the average mini-clinical evaluation scores, the percentage of logbook

quota completed, the clinical case presentation score, and the blinded patient case score.

Data analysis

The data was analysed using STATA version 15. Descriptive statistics were used to assess the characteristics of the continuous workplace-based assessment components i.e. means, proportions, standard deviation, median and maximum and minimum scores. WPBA scores for each clinical encounter were plotted to track student progress. Quantile regression of the median scores of the mini-CEX encounters were used to test for a linear trend in weekly scores, adjusting for the repeated measures within students. A Spearman's rank-order correlation was run to assess the relationship between the various assessment components. This was illustrated using correlation matrices. The closer the correlation coefficient is to the value of 1, the better the correlation. One-sample *t*-tests were used to determine whether there was a statistically significant mean difference between various scores i.e. when the *p*-value is less than or equal to 0.05. Where outliers were detected, they were included in the analysis and the assumption of normality was tested by Shapiro-Wilk's test ($p > 0.05$).

In addition to statistical analysis, two qualitative approaches were used to assess the reliability and validity of the WPBA practices as discussed previously in section 2.1. Driessen *et al.*'s (2005) principles of credibility and dependability were applied to assess the reliability of the assessment, and validity was evaluated using a combined framework based on Kane (2013) and Royal (2017).

4.8. Ethical considerations

The project was submitted to the Stellenbosch University Health Research Ethics Committee (HREC) for Ethics Approval (Ethics Reference #: X18/02/002- Appendix F). A request for a waiver of individual informed consent was made as this was a retrospective review of routinely collected data with minimal risk. The Research Committee of the Faculty of Dentistry noted and approved the request to access the data for the purposes of the study and consent was also obtained from the Registrar of the University of the Western Cape as the custodian of this data (Reference number: UWCRP070318NM- Appendix G)

Data was extracted from student records. A risk to participants is thus a breach of confidentiality. To minimize this, several steps were undertaken:

- An arbitrary anonymization key was developed to remove any identifying information of the students.
- A list, linking the identifiable data to the unique study number was created and stored separately in case there was a need to re-check data entered into the database for inspection purposes.
- Data was entered in electronic databases that were password protected.
- Data was only communicated in anonymous form for analysis purposes.
- Hard copies of the data were stored in a locked cabinet.
- All identifiable information has been omitted from the final report.

4.9. Positioning of the researcher

The researcher plays an important role in the success of the research (Unluer, 2012). Being an insider researcher could help with better understanding of the phenomena and context being studied. This could however also lead to a loss of objectivity.

My position in the faculty is lecturer and clinical supervisor in the Department Paediatric Dentistry. This could therefore positively impact on my role as researcher of this study as I am intimately involved in the assessment processes and teaching (clinical and theoretical) in the department. I have the advantage of experiencing the implementation of the continuous WPBA practices first-hand, which includes feedback from staff and students. A disadvantage is the danger of making assumptions based on anecdotal observations or pre-held ideas. This was consciously addressed with my supervisor throughout the process.

CHAPTER 5: RESULTS

The study population consisted of 57 and 58 students in the 2016 and 2017 cohorts respectively i.e. a total of 115 complete records. A total of 60 students were excluded from the study as the full complement of hard copy records was not available. This included 36 records for 2016 and 24 records for 2017.

5.1. Summary of assessment results

A summary of the study cohort assessment data is presented in Table 5.1.1 as mean (%) and minimum and maximum scores for each assessment component (with standard deviations). A schematic presentation of the same data is represented in Figure 5.1.1. as a box-and-whisker plot, depicting the central tendency and spread as median, with upper and lower quartiles. This visual presentation of the data additionally illustrates the outliers in the data set.

As discussed before in section 3.1 and 3.2, the average mini-clinical evaluation score (Ave mini-CEX) represents the final mini-clinical evaluation mark that was obtained at the end of the year after all clinical encounters were taken into consideration. The average case presentation mark represents the average of the two case presentations during the year. The logbook quota score represents the weighted percentage after completion of required procedures. The scores for the blinded patient case (BPC) refer to the single clinical assessment at the end of the module in the final year. The combined continuous WPBA score includes all these assessment components.

Table 5.1.1: Summary of assessment data for the cohort (n=115)

Assessment component	Average mini-CEX scores (%) [with SD]	Ave case presentations (%) [with SD]	Logbook quota (%) [with SD]	Blinded patient case (%) [with SD]	Combined Continuous WPBA score (%) [with SD]
Mean	58.9 [4.3]	65.3 [6.4]	63.2 [11.6]	67.8 [8.2]	61.9 [4.9]
Minimum score	50	36.5	35	45	49.3
Maximum score	67.5	81.2	90	83	74.9

The highest scores were recorded for the BPC component at the end of the semester with the worst scores recorded in the daily clinical assessments (mini-CEX). The average mini-CEX scores demonstrate the lowest standard deviation though, with consistent scores being allocated to students over multiple opportunities by various assessors.

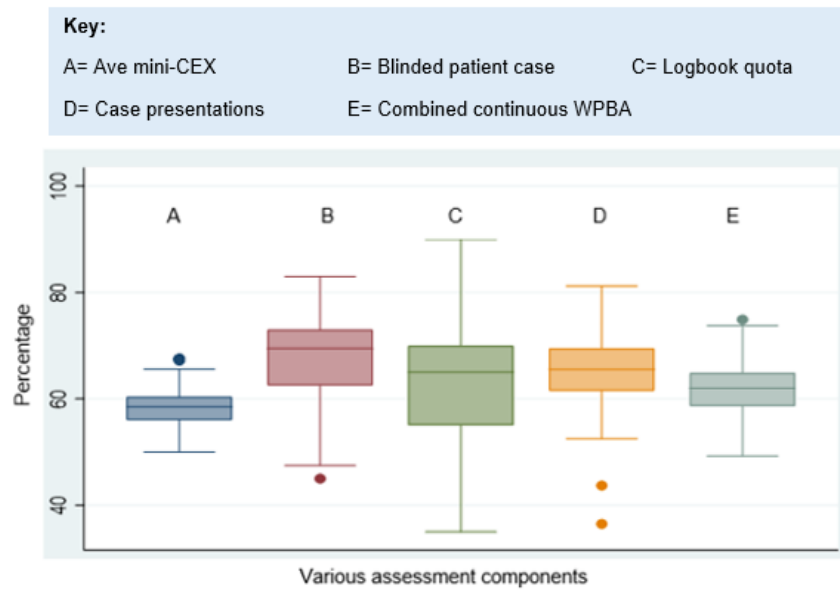


Figure 5.1.1: Summary of assessment data

5.2. Correlations between the various components of the continuous workplace-based assessment

Positive correlations were found between most of the assessment components as depicted in Figure 5.2.1 and Table 5.2.1.

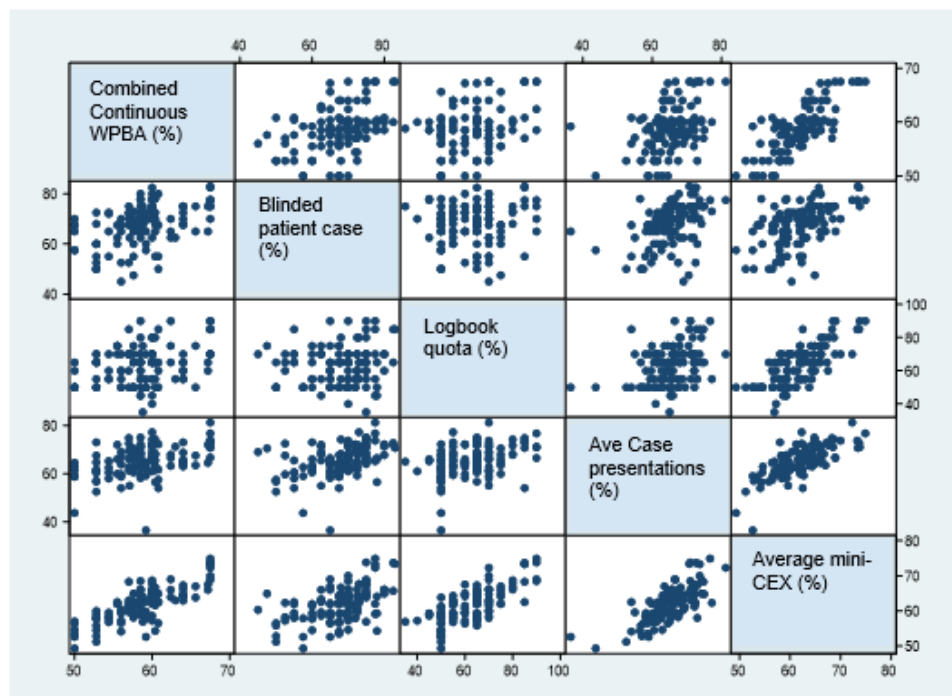


Figure 5.2.1: Scatterplot matrix depicting correlations between various assessment components for the combined 2016-2017 cohort

Table 5.2.1: Spearman correlation coefficients between various assessment components (correlation row 1, *p*-value row 2)

	Ave mini-clinical evaluation	Cases ave	Logbook Quota	Blinded patient case	Combined continuous WPBA
Average mini-clinical evaluation	1.000				
Case presentation average	0.4114 0.0000	1.0000	0.3710 0.0000	0.4848 0.0000	
Logbook quota (procedures)	0.2550 0.0060		1.0000	0.0973 0.3011	
Combined continuous WPBA	0.7218 0.0000	0.7229 0.0000	0.7192 0.0000	0.5289 0.0000	1.0000
Blinded patient case	0.4704 0.0000			1.0000	

A moderate positive correlation was found between the average mini-CEX and the BPC ($r_s = 0.47$, $p < 0.0001$) which was statistically significant (Table 5.2.1; Figure 5.2.2). Participants performed better in the scores obtained for BPC (Table 5.1.1), 67.8% (8.2) as opposed to the scores obtained for their average mini-clinical evaluation, 58.9% (4.3), a statistically significant difference of -8.82% (95% CI, -10.16 to -7.48), $p < 0.0001$.

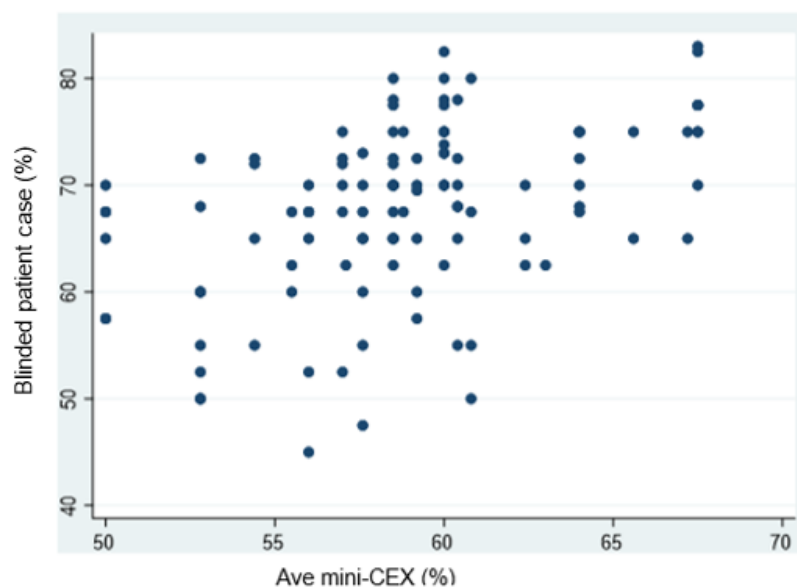


Figure 5.2.2: Correlation between average mini-clinical evaluation and the BPC

The average of the case presentations was moderately correlated with the blinded patient case scores ($r_s=0.52$, $p < 0.001$) which was statistically significant (Table 5.2.1; Figure 5.2.3). Participants performed better in the BPC (Table 5.1.1), 67.8% (8.2) as opposed to the average scores obtained for the case presentations, 65.3% (6.4), a statistically significant difference of -2.39% (95% CI, -3.86 to -0.92), $p = 0.0017$.

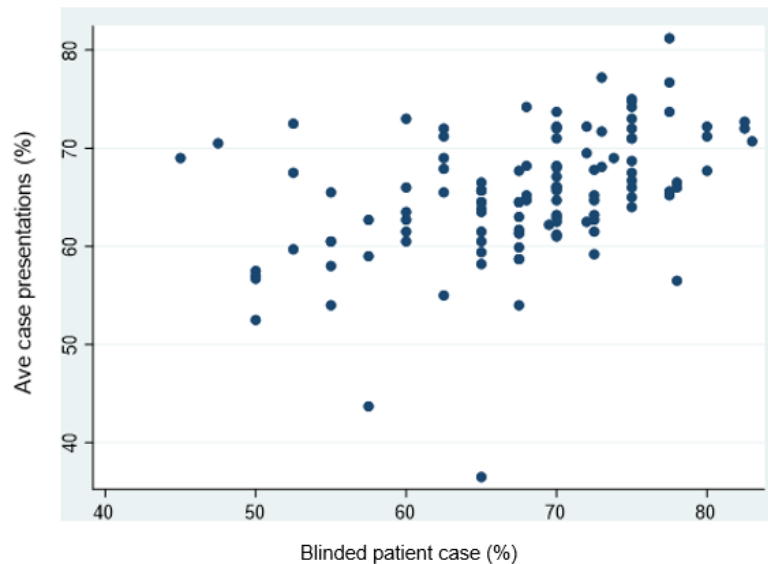


Figure 5.2.3: Correlation between the case presentations (average) and the BPC

The average mini-CEX scores had a weak positive correlation with the logbook quota ($r_s = 0.25$). This was statistically significant ($p = 0.006$). (Figure 5.2.4).

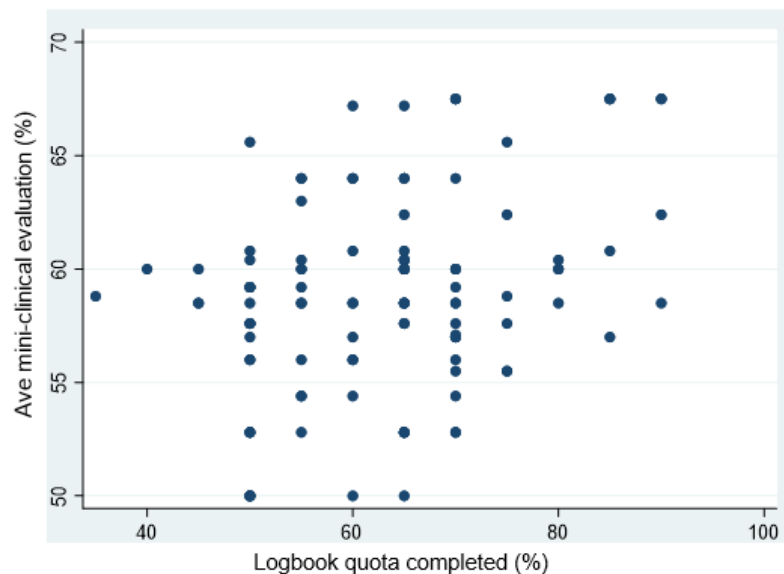


Figure 5.2.4: Correlation between average mini-clinical evaluation and percentage of logbook quotas completed

Both the average mini-CEX scores and logbook quotas are expressed as a percentage. The weak correlation between the two components indicates that the average clinical mark obtained in the clinics at the end of the year did not necessarily reflect the amount of work or number of procedures i.e. logbook quota that was completed.

The combined continuous WPBA moderately correlated with the blinded patient case ($r_s=0.52$), $p < 0.001$ which was statistically significant (Table 5.2.1; Figure 5.2.5). Participants performed better in the single clinical assessment, 67.8 (8.2) (Table 5.1.1) compared to their combined continuous WPBA, 61.9 (4.9). A statistically significant difference was noted, -5.74 (95% CI, -7.03 to -4.45), $t(114) = -8.79$, $p < 0.0001$.

The combined continuous WPBA scores were also strongly correlated with the average mini-CEX score ($r_s=0.72$, $p < 0.001$) as well as quota and cases ($r_s=0.72$, $p < 0.001$). All of these findings were statistically significant (Table 5.2.1).

It has to be noted as part of the interpretation of this part of the analysis, that the single BPC score is included in, and contributes 20% to the combined continuous WPBA score.

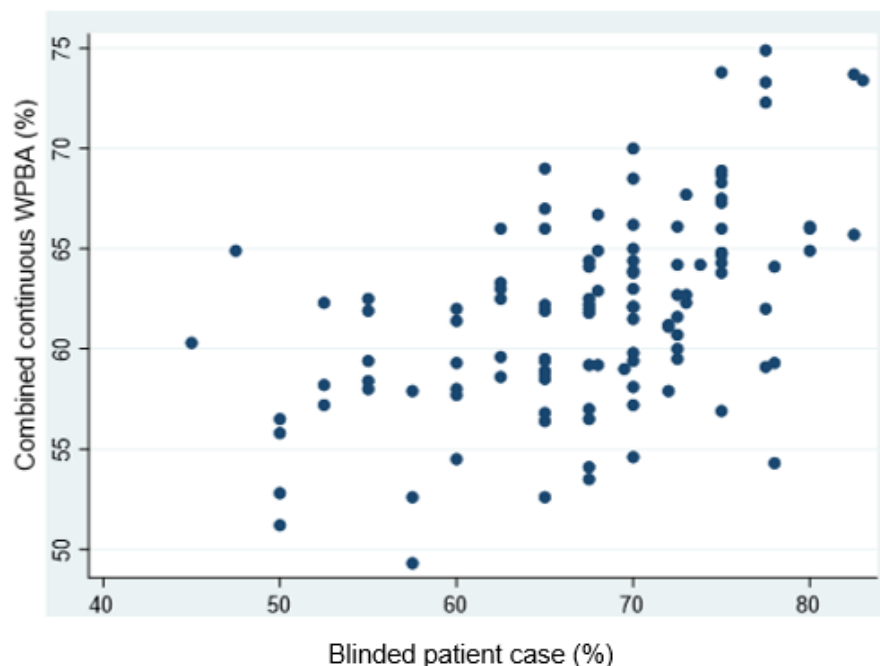


Figure 5.2.5: Correlation between the combined continuous WPBA scores and the blinded patient case

5.3. Identification of struggling students

Due to the low failure rate, predictive values could not be calculated. Table 5.3.1 represents an excerpt of students who have failed any of the assessment tasks.

Only one student (B) obtained less than 50% i.e. 49.3% for the combined continuous WPBA score. This student achieved borderline scores for both the average mini CEX and logbook quota, whilst failing the case presentation. Based on Faculty's assessment guidelines, the mark was condoned to 50% to allow the student entrance to the final examination as 50% was obtained in the average mini-CEX and quota (Refer to section 3.1).

Two students (A and B) failed the case presentations. Whilst both their average mini-CEX and logbook quota marks were low, the blinded patient case scores were their best assessment scores.

Two students (C and D) failed the end of module blinded patient case despite consistent performances across the other components.

Five students (E to I) could not obtain the minimal procedural quota of 50%, but performed consistently well across the other assessment tasks; with their blinded patient case their highest scores.

No student obtained less than 50% in more than two clinical components.

Table 5.3.1: Summary of marks of students who failed a component of the assessment tasks

Student	Ave mini-CEX% [50%]	Logbook Quota (%) [20%]	Case presentation (%) [10%]	BPC (%) [20%]	Combined continuous WPBA (%) [Final Clinical Paediatric mark]
A	59.2	50	36.5	65	52.6
B	50	50	43.7	57.5	49.3
C	56	70	69	45	60.3
D	57.6	75	70.5	47.5	64.9
E	58.5	45	65.2	77.5	59.1
F	60	45	65.5	62.5	58.6
G	60	40	61	70	57.2
H	58.8	35	65	75	56.9
I	58.5	45	66	78	59.3

5.4. Reliability of WPBA

5.4.1. Quantitative Approach

The available retrospective data did not lend itself to a formal reliability analysis since there was no overlap across students or encounters. A prospective study with an experimental design across students, supervisors, time and patients is recommended.

A sub-analysis of mini-CEX performances and examiner scoring distribution was however done to describe performance variation over time.

Mini-CEX performance over time

Each student was assessed over a period of approximately 30 weeks. Not all students had the same number of mini-CEX encounters. Each student treated a minimum of one case per week and a maximum of two cases per week in the workplace during the academic year.

Figures 5.4.1.1 and 5.4.1.2 depict the allocated scores for the mini-CEX encounters per week during 2016 and 2017 respectively. The y-axis represents the scores allocated according to the clinical rubric i.e. categories 1 to 5 as previously explained in Chapter 3 (Appendix C).

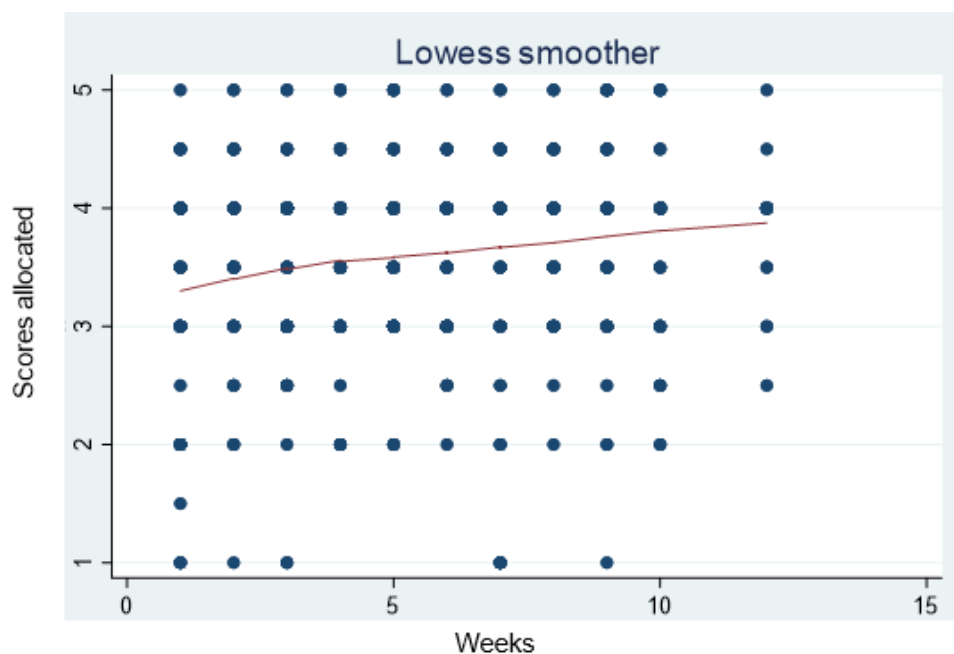


Figure 5.4.1.1: Allocated scores for mini-CEX encounters over time (2016)

When testing for a linear time trend in the weekly scores and adjusting for the repeated measures within using quantile regression of the median scores, a significant positive improvement over time was noted ($p < 0.001$, slope = 0.0625 per week) in the 2016 group of students. On average, the score increased by 0.625 over the 10-week period which roughly equates to a 16% increase in scores from the end of the first semester 2016 to the end of the 2nd semester 2016. This may have been attributed to the introduction of the new clinical rubric in the middle of 2016. No improvement in weekly clinical scores could be detected in the 2017 cohort using the same clinical rubric (Figure 5.4.1.2). Possible reasons for these results will be addressed in the discussion section.

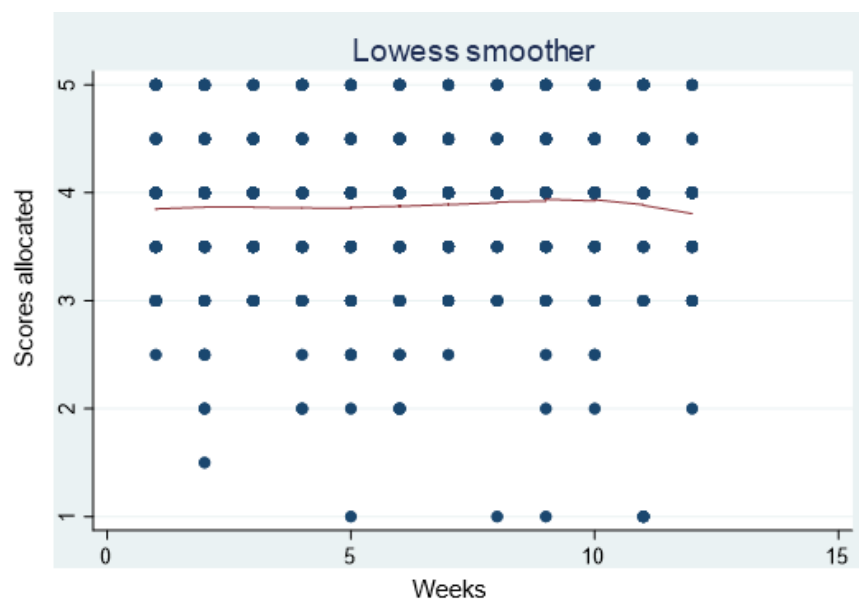


Figure 5.4.1.2: Allocated scores for mini-CEX encounters over time (2017)

Scoring distribution of examiners

To determine the scoring distribution of individual clinical supervisors, a sub-analysis of scores allocated by individual examiners was done with available data on diagnostic encounters in the workplace (Table 5.4.1.1).

Table 5.4.1.1: Scoring profiles of clinical supervisors based on scores given according to rubric (scores 1 to 5) for diagnostic encounters

Examiner	Number of encounters per scores given				
	1 (=20%) <i>clear fail</i>	2 (=40%) <i>fail</i>	3 (=50%) <i>pass</i>	4 (=60%) <i>clear pass</i>	5 (=75%) <i>distinction</i>
1	2	3	4	22	7
2	2	0	12	8	0
3	1	4	15	64	0
4	0	1	18	90	11
5	1	3	12	48	6
6	1	4	20	48	5
7	2	9	16	26	1
8	0	4	41	47	20
9	0	4	14	22	10
10	0	5	21	45	6
11	0	1	20	83	10
12	0	0	2	40	2
TOTAL	9 (1%)	38 (4%)	195 (23%)	543 (63%)	78 (9%)

The most frequent score given was 4 (63% of cases) which is equivalent to a score of 60% and a clear pass (Appendix C). The clinical and theoretical descriptors of this specific category are provided in Table 5.4.1.2.

Table 5.4.1.2: Interpretation of a score of 4

Clinical	Theory
<ul style="list-style-type: none"> Missed some things regarding history but could answer when prompted Treatment plan acceptable. Needs only minor revision. Very limited guidance needed with procedure Good quality restoration (good contacts/ marginal adaptation)-- minor adjustment needed 	<ul style="list-style-type: none"> Sound knowledge (better than average) <ul style="list-style-type: none"> Good motivation but doesn't cover all the possible treatment options Competent Able to justify material selection Good insight demonstrated but can improve Theoretical knowledge good

Further comparison of these profiles showed significant differences in scoring amongst individual examiners. Given the marginal results of the supervisors, reliability is expected to be moderate.

5.4.2. Qualitative approach

Clinical supervisors have to evaluate and judge critical thinking and the quality of student responses when students motivate treatment choices and apply theory to the clinical context as part of the mini-CEX evaluations and case presentations. As previously discussed in Chapter 2, this is a subjective interpretation of the student's abilities with assessment information also presented as qualitative data where the principles of credibility and dependability can be applied.

Using this qualitative analysis strategy as proposed by Driessen *et al.*, (2005), principles of credibility and dependability were applied to the Paediatric Dentistry assessment practices based on assessment records and information contained in the study guides (Refer to section 2.2). A summary of the evaluation is described in Table 5.4.2.1.

Table 5.4.2.1: Evaluation of reliability of the assessment approach using qualitative strategies (Driessen *et al.*, 2005)

	Strategies	Application to Paediatric Dentistry	In- text reference
CREDIBILITY OF ASSESSMENT	Triangulation (combining information from different sources)	✓ Varied assessment used (mini-CEX, logbook quota, case presentations, blinded patient case)	Sections 2.1, 3.1 and 3.2
	Prolonged engagement over time	✓ Multiple assessments over time ✓ Multiple examiners ✓ Multiple and varied patient cases	Sections 2.1, 2.2, 3.1, 3.2 and 5.4.1
	Member checking	✓ Student feedback ✓ Student involvement in assessment process	Sections 2.1, 3.1 and 3.2
DEPENDABILITY	Quality assurance	✓ Rubrics with explicit criteria ✓ All staff attended training ✓ All assessment data is well documented (with additional comments) ✓ Regular evaluation and adaptation of assessment practices based on staff and student feedback	Appendix C Sections 2.1, 3.1 and 3.2

This qualitative analysis suggests that the continuous formative WPBA results are reliable.

5.5. Validity of WPBA

In order to structure an argument for validity, a combination of two frameworks (see chapter 2, section 2.2) was used that most accurately represented the WPBA practices of the department i.e. Kane's four inferences (2013) and Royal's four tenets (2017). The findings of the application of these frameworks are summarised in Table 5.5.1.

Table 5.5.1: Validity argument for WPBA in Paediatric Dentistry

Tenets/ inferences	Application to Paediatric Dentistry	In- text reference
SCORING	<ul style="list-style-type: none"> ✓ Use of a detailed clinical rubric ✓ Staff trained to use rubric ✓ Evaluates both clinical and theoretical aspects for each patient case 	Sections 2.2, 3.1 and 3.2
GENERALISATION/ VALIDITY EVIDENCE	<ul style="list-style-type: none"> ✓ Multiple assessments ✓ Multiple examiners ✓ Varied patient cases 	Sections 2.1, 2.2, 3.1, 3.2 and 5.4.1
IMPLICATIONS	<ul style="list-style-type: none"> ✓ Correlation of scores between WPBA components ✓ WPBA as an indicator of clinical competence ✓ Value of end of module clinical assessment 	Section 5.2

**Adapted from Royal (2017) and Kane (2013)

This analysis suggests that the continuous formative WPBA results are valid.

CHAPTER 6: DISCUSSION

The purpose of this study was to determine the extent to which departmental WPBA scores could be used as an indication of clinical competence of final year undergraduate dental students. Correlation between the various assessment components was determined to ascertain whether meaningful conclusions could be drawn from the assessment scores that would help inform decisions regarding current departmental assessment practices.

No studies in the literature could directly be compared with the combination of assessment practices employed by the Department of Paediatric Dentistry. Where applicable, the individual components will therefore be discussed in relation to the literature.

6.1. The impact of continuous formative assessment

The continuous WPBA tasks in our study are formative in nature and improvement in performance may thus be expected over time (Norcini & Burch, 2007). This assessment for learning where feedback is provided is essential for student growth and development (Konopasek *et al.*, 2016).

In the present study, an improvement in mini-CEX scores were however only demonstrated in the 2016 class. This may not be a true reflection of improvement as the difference could be attributed to the introduction of a new scoring rubric at the start of the 2nd semester in 2016 as previously discussed in Chapter 3. Scores remained consistent after the introduction of the new rubric for both the 2016 and 2017 cohorts.

Consistent performance may have been influenced by the study cohort characteristics. As final year students who are close to graduating, they should possess or be close to possessing the necessary skills to successfully enter the world of work. Thus, an improvement might not be evident as most have already achieved the expected level of competency. The variation in clinical sessions could further play a role in these findings. Patients, supervisors and clinical procedures differ for each session. As each of these variables is different, it is difficult to assess a definite improvement from one episode to the next. If the same procedure is repeated under supervision of the same supervisor, it may be easier to detect whether an improvement has taken place.

The scoring scale of the assessment rubric may also influence this 'lack' of improvement shown in our study. Examiners broadly assessed students according to 5 categories, where most students would be expected to achieve a pass (3) or a clear pass (4). The scoring system may thus not be sensitive enough in identifying small improvements.

6.2. Formative results as predictors for summative performance

In our study, the average scores obtained for the end-of module blinded patient case were statistically significantly higher than the average scores obtained in all the other assessment components during the module. (Table 5.1.1, Figure 5.1.1). This is similar to findings reported by Carrillo-de-la-Pena *et al.*, (2009) where success in formative assessment tests were associated with better summative marks. A total of 548 students enrolled for three different health sciences degrees from four Spanish universities participated in that particular study. Anziani *et al.*, (2008) and Riaz *et al.*, (2015) further supported this positive contribution of formative assessment to summative results.

Krasne *et al.*, (2006) reported the use of formative assessments as effective predictive tools of summative performance in medical schools. In our study, none of the continuous workplace-based assessment components could predict the end-of-module clinical examination score though.

It has to be noted however that our end of module assessment comprises of a single patient case. These clinical cases are not standardised and vary in terms of presenting problems and complexity. The reliability of these results therefore have to be questioned. This score may thus not be a true reflection of competence when compared to the other assessments that take place over a period of time. However, the fact that the scores obtained for the end-of-module blinded patient case were better even though a more detailed analytic rubric was used, allowing for a more structured assessment process, can also be seen as evidence of improved performance.

The continuous assessment results could be a better indication of student performance and could be considered when making a judgement call regarding whether or not a student should pass in our setting; especially in view of the feasibility concerns discussed in Chapter 3. Harris *et al.*, (2017) supports this view of multiple assessments being a more accurate indicator of whether a student is fit to progress to the next level.

6.3. The value of multiple and varied assessments

Clinical practice requires a range of characteristics, as no single method of assessment is likely to provide enough data to make a valid and reliable judgement on competency (Norcini & Burch 2007). Assessment methods should therefore be fit for purpose with an understanding of the information it can provide. Individual assessment methods have strengths and weaknesses and issues of reliability, validity and feasibility should be considered. Multiple and varied assessment tasks further provide students with the opportunity to showcase different competencies and strengths, making the assessment more fair (Downing & Yudkowsky, 2009; UWC assessment policy, 2017).

Whereas most WPBA methods can be used for formative assessment on their own, multimodal assessment across a time period is recommended for summative decisions (Schuwirth & van der Vleuten, 2010). Clinical competence assessment in the Paediatric Dentistry programme is levelled at the 'does' level of Miller's pyramid (Miller, 1990; Wass *et al.*, 2001). Various WPBA methods are used over time to assess performance.

Because clinical skills are essential for a dentist, it is important that competence when performing clinical procedures, is assessed. A logbook is used to monitor the number and type of clinical procedures students have completed. Whilst Dahllöf *et al.* (2004) reported on the use of logbooks for reflection purposes and Anziani *et al.* (2008) used logbooks to compare formative and summative assessment scores, no studies have examined the correlation between clinical performance and number of procedures completed. The average mini-CEX score in this study showed only a weak positive correlation with the logbook quota (Figure 5.2.4). Students that obtained good clinical marks have thus not necessarily completed a higher percentage of the required quota. The majority of students who failed one assessment task, did so for this particular component. As previously discussed in Chapter 3, not meeting the procedural quota could be attributed to various factors which are beyond the student's control. Even though none of these students failed the paediatric module, the contribution of the quota to the final mark and inferences made regarding competency, have to be reconsidered, especially in light of the fact that some departments in the faculty use logbook quotas as the sole means of assessing clinical competence. The results of this study show that the number of procedures does not equate to clinical competence and that a more holistic view of student performance should be considered.

Presentation of patient cases seen in the clinical area allows students to appraise their chosen treatment options and defend their decisions with appropriate motivation from

the literature. This is different from the real-time mini-CEX assessments as students have time to reflect on their behaviour and choices. This helps to hone critical thinking and problem-solving skills and improve understanding of complex issues through integration of concepts (Popil, 2011; McDade, 1995). Two students failed this component (Table 5.3.1). Both students' mini-CEX assessment performance was relatively poor and they only just passed the procedural quota assessment component. The mini-CEX and case presentations both assess integration and articulation of knowledge and skills; and could identify the struggling students in this cohort who needed further support.

The findings of this study suggest that the combined WPBA components provide a more accurate reflection of student competence than a single assessment task. This is in line with the principles of programmatic assessment which states that more than one component should be evaluated to obtain a more holistic view of student competence (van der Vleuten *et al.*, 2015). The combined continuous WPBA score includes all clinical work, case studies and quota of procedures. This is therefore a true representation of the workplace as integration of theoretical knowledge and the application thereof to the clinical context is assessed. This integrated assessment as discussed in Chapter 1, is important to determine clinical competence.

6.4. The case for reliability

Jonsson & Svingby (2007) suggested that multiple examiners and rubrics with explicit criteria strengthens reliability, especially if criteria are strictly adhered to. In the Paediatric Dentistry department all departmental staff have been trained to use the rubrics and emphasis was placed on identifying specific criteria to substantiate the mark that is allocated. Regular reinforcement is nonetheless essential to ensure that staff are reminded of the process and remain focussed on the purpose of the assessment. This would help to improve accuracy when completing the rubric.

The learning outcomes and the purpose of the assessment were used when designing the rubric. As recommended by Gibson and Shaw (2010), specific aspects/ characteristics that are linked to the learning outcomes and that need to be evaluated were identified and incorporated into the rubric. Rubrics that are specifically designed for a particular clinical context also enhances its reliability (Jonsson & Svingby, 2007). This is therefore applicable to the clinical rubric which is specifically designed for the Paediatric Dentistry context.

The use of a holistic rubric for the daily mini-CEX encounters means that the scores that students can achieve fall within a certain range which is not very specific. The score distribution for the blinded patient case had a wider range and demonstrated greater discrimination potential between the high and low scorers. On the other hand, if the rubric scale is too short, identifying small differences between students is more difficult (Perlman, 2003). This was the case with the mini-CEX evaluations. The small standard deviation for the mini-CEX scores generated from the current study nonetheless shows that the data is clustered closely around the mean, which may be an indication of reliability (Table 5.1.1 & Figure 5.1.1).

With the once-off clinical assessment, students can either perform very well or very poorly depending on the type of patient case on that particular day. This is reflected in Table 5.3.1 where two students failed the blinded patient case despite consistent performances across the other components. This could either be attributed to the complexity of the patient case or the students' inability to handle high-pressure situations. This has been confirmed by Wass *et al.*, (2001) who highlighted the fact that these conditions could influence the reliability of the assessment

Cook and Beckman (2009) compared a five-point mini-CEX scoring scale with a nine-point scale. They found that even though the nine-point scale appeared to be more accurate, the inter-rater reliability was similar for the two scales. Similarly, the ten-point scale used for the blinded patient case is likely to be more accurate with regards to pinpointing actual student performance but if the purpose of the assessment is simply to ascertain whether a student is competent or not, a five-point scale as with the mini-CEX is sufficient. A trade-off therefore has to be made between accuracy and efficiency. In a busy clinic, a holistic rubric is less time-consuming and more practical, yet it still provides enough information to enable the student to improve on future attempts.

The criteria delineated in the holistic clinical rubric portrays more than one description for each level. The clinical rubric has been shown to be reproducible. The fact that staff and students often agree on the same score, attests to the clarity of the assessment criteria as stipulated on the rubric. Chances for misinterpretation are low, especially for the clinical component. Overt performance where actual skills are assessed is easier to measure (Andreatta *et al.*, 2009). However, the theoretical component that underpins the clinical practice is more open to subjective interpretation because covert performance like clinical reasoning is assessed (Andreatta *et al.*, 2009). The fact that examiner subjectivity does come into play with the mini-CEX evaluations makes it vulnerable to

bias. On the other hand, as posited by Pelgrim *et al.*, (2011) multiple assessors counter the effect of subjectivity of the assessor.

Knowledge and experience of the clinical supervisor (Govaerts *et al.*, 2011) play a vital role in the extent to which staff are able to interpret the quality of student responses and grade them fairly. Expertise of the assessor influences how information regarding student performance is processed (Govaerts *et al.*, 2013). Experienced supervisors tend to look at the situation more holistically i.e. by linking specific tasks to behaviours and outcomes (Govaerts *et al.*, 2013). Moreover, it should be borne in mind that factors like inherent personality traits, beliefs and opinions of staff can also influence how students are scored (Tziner *et al.*, 2005). Scoring can therefore never really be “objective” as there is “no single true score” yet, it should be noted that all perspectives are valid (Govaerts & van der Vleuten, 2013, p.1170). The fact that the average mark for the two components gets recorded does however level the playing field to some extent as any discrepancies resulting from subjective interpretation on the part of the clinical supervisor is mitigated somewhat.

A strong correlation between the combined continuous WPBA score and the individual WPBA components indicates good reliability of the assessment (Table 5.2.1). Students are assessed by multiple supervisors who rotate through the clinic at different times and treat a broad spectrum of patients with varied needs. Reliability and validity is thus improved as students are assessed by more than one supervisor on multiple occasions over a period of time. This is in agreement with sentiments expressed in the literature (van der Vleuten & Schuwirth, 2005; Downing, 2004). Pelgrim *et al.*, (2011) found that an acceptable level of reliability can be achieved using a minimum of ten encounters. The number of encounters included in the present study is more than double this. As supported by van der Vleuten & Schuwirth (2005), the larger sample accounts for variance and minimises errors, thereby improving reliability. The outcome of the assessment can therefore be considered reliable.

Quality assurance is an indicator of dependability which is another strategy suggested by Driessen *et al.*, (2005) to confirm reliability. Programmatic assessment which has input from all parties involved in the assessment process can provide important information on the quality of the assessment processes (van der Vleuten *et al.*, 2015). This is extremely important to ensure that assessment processes are as sound and fair as possible. The assessment processes in the Department of Paediatric Dentistry are continually being revised based on staff and student feedback to ensure fairness,

transparency and improved adherence to the principles of sound assessment. Student involvement in the assessment process is said to enhance transparency and the “quality of the assessment” (Falchikov, 2003, p.106).

Based on Driessen *et al.*'s (2005) qualitative approach to reliability, the continuous formative WPBA practices in Paediatric Dentistry could be viewed as credible and dependable, and therefore reliable (Table 5.4.2.1).

6.5. A case for validity

According to Downing (2004), reliability of an assessment is the main indication of its validity. The proven reliability of the WPBA in Paediatric Dentistry is therefore the first source of evidence in favour of validity of the departmental assessment practices. Content validity suggests the entire course content should be covered (Schuwirth & van der Vleuten, 2010) and that competencies should be aligned with the learning objectives (Coderre *et al.*, 2009). The importance of constructive alignment was also highlighted by Biggs (1999). In the module being evaluated, constructive alignment exists between assessment methods employed and the learning outcomes as evidenced in the module descriptors and study guides. As referred to in section 3.2, the latter sources were used when compiling the clinical rubric. Content is specific to the discipline and content validity is therefore enhanced. This is supported by Durning *et al.*, (2002).

Construct validity implies that an assessment tool should be able to discriminate between the high and low scorers (Schuwirth & van der Vleuten, 2010). It would seem as though the clinical rubric does not discriminate between very good and very poor students. The tendency for departmental staff (.e. part-time and permanent) to cluster their scores around a '4' (Table 5.4.1.1) does however suggest that supervisors find students to be competent. This score correlates well with the average mini-CEX scores over the entire year (Table 5.1.1). The latter all lie in the vicinity of 60%. A flaw in the rubric is that the upper score limit is recorded as 75%+ which is not very specific and could potentially disadvantage students at the upper end of the spectrum. Consideration could be given to increasing the weighting for this upper category or adding an additional category to cater to the exceptional students.

Al Ansari *et al.*, (2011) demonstrated evidence of construct validity over time with the mini-CEX evaluations. Over a shorter period of observation, differences in performances were less significant compared to when performance was assessed over more than one year of residency training. Thus, tracking students over their fourth and fifth years of

study in the Paediatric Dentistry module is more likely to show a greater improvement over time compared with only evaluating performance over the final year of study where students are generally expected have the necessary expertise and are basically just refining their skills.

Where the blinded patient case is concerned, each student is examined on a different patient case and is therefore not standardised. This means that the complexity of the cases may vary between students, which is not ideal. This therefore negatively affects the fairness and validity of the blinded patient case component and provides motivation for why consideration should be given to removing the once-off, blinded patient case from the assessment.

According to Andreatta *et al.* (2009), validity refers to whether decisions based on assessment data within a particular context can be trusted. Based on the combined framework of Kane (2013) and Royal (2017), evidence of validity in the departmental WPBA have been highlighted (Table 5.5.1). All individual assessment components are taken into consideration when making a decision on whether or not a student is competent i.e. programmatic assessment practices are followed. This information richness adds to the credibility of pass/ fail decisions that are taken (van der Vleuten *et al.*, 2015). Validity is thus enhanced as a more complete representation of a student's abilities can be obtained (Schuwirth & van der Vleuten, 2010).

The construct being assessed, skills, knowledge and application of knowledge to a particular patient case, is clearly defined in Paediatric Dentistry with the measurement tool, i.e. the detailed clinical rubric, measuring what it is supposed to measure in an actual clinical setting. The overall workplace-based assessment practices can therefore be considered valid within this context.

6.6. Limitations of this study

This study looked at retrospective assessment data, which was incomplete and influenced the ability to perform statistical analysis. This in turn had an impact on the conclusions that could be drawn.

Even though it was originally confirmed by the statistician that reliability could be determined, after data collection, it was advised that this was not possible due to the nominal nature of the data. As retrospective data was used in the present study, inter- and intra-rater reliability (i.e. consistency over multiple assessments) could not be

determined. This also means that it was not possible to assess “rater error” and how it impacts on the score (Cook and Beckman, 2009, p 656). Multiple variables as mentioned previously could also impact on the scoring. More than one examiner should assess a case to determine the inter-rater reliability. Due to staff numbers in our context, this is not feasible and correlation of scores between raters for the same patient case was not possible. Only data that was available could be used, thus the number of incomplete records that were excluded, compounded the problem.

The introduction of a different clinical rubric in the middle of 2016, is a further limitation although it assessed a similar construct. Furthermore, this study is a single-centre review and findings may not be transferable to other settings, although other resource-poor settings may share similar challenges. Being an insider researcher could also be considered an additional limitation.

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS

The aim of the study was to determine whether the current WPBA results in the Department of Paediatric Dentistry at the University of the Western Cape could be used as an indication of clinical competence of final year students.

The continuous formative WPBA practices in the Department of Paediatric Dentistry have proven to be feasible as it is integrated in the daily routine patient care provided in the paediatric dental clinics. The continuous formative WPBA scores further seem to be an indication of clinical competence of final year dental students at UWC. On the other hand, this study showed that the addition of the single BPC could be reconsidered due to feasibility and reliability concerns.

Further prospective research with an experimental study design across students, supervisors, time and patients would be required to increase our confidence in the inferences made based on current assessment results. Assessing students over the entire Paediatric Dentistry course i.e. over the 4th and 5th years of study would provide more accurate information regarding student progress and perhaps also indicate where intervention might be needed. This therefore provides a potential opportunity for future research. In addition, as the use of the clinical rubric was not reinforced in 2017, the possibility that staff did not strictly adhere to the rubric and perhaps lapsed into a measure of complacency, cannot be discounted. The effect that faculty development initiatives and staff training could have on student performance should be investigated further especially regarding the accurate utilization of rubrics by staff.

Where workplace-based assessment provides the opportunity to assess students on all aspects of patient management in an authentic clinical environment, the complexity of the workplace-based setting poses a challenge where assessment practices are concerned. This is especially true when having to make a judgement call regarding a student's clinical competence. This study suggests that holistic evaluation of student performance is essential when making such a judgement call. A balance must be struck between retaining good clinical practice yet, making sure assessment practices are feasible, fair and more importantly, reliable and valid. The final decision on the competence of a student should be made on the basis of professional judgements which are supported by evidence that is both defensible and credible.

REFERENCES

- Adamson, K., Gubrud-Howe, P., Sideras, S. & Lasater, K. (2012). Assessing the reliability, validity, and use of the Lasater clinical judgement rubric: three approaches. *Journal of Nursing Education*, 51 (2): 66- 73.
- Al Ansari, A., Kauser, S. & Donnon T. (2013). The construct and criterion validity of the Mini-CEX: A meta-analysis of the published research. *Academic Medicine*, 88 (3): 1-8.
- Albanese, M.A. (2000). Challenges in using rater judgements in medical education. *Journal of Evaluation in Clinical Practice*, 6 (3): 305- 319.
- Andreatta, P.B. & Gruppen, L.D. (2009). Conceptualising and classifying validity evidence for simulation. *Medical Education* 43: 1028–1035
- Anziani, H., Durham, J. & Moore, U. (2008). The relationship between formative and summative assessment of undergraduates in oral surgery. *European Journal of Dental Education*, 12: 233- 238.
- Azzi, A.J., Ramnanan, C.J., Smith, J., Dionne, E. & Jalali, A. (2015). To quiz or not to quiz: Formative tests help detect students at risk of failing the clinical anatomy course. *Anatomical Sciences Education*, 8 (5): 413- 420.
- Beard, J. (2011). Workplace-based assessment: The need for continued evaluation and refinement. *The Surgeon*, 9: S12- S13.
- Biggs, J. (1999). What the student does: teaching for enhanced learning. *Higher Education Research & Development*, 18 (1): 57- 75.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5 (1).
- Boulet, J.R., McKinley, D.W., Norcinic, J.J. & Whelan, G.P. (2002). Assessing the comparability of standardized patient and physician evaluations of clinical skills. *Advances in Health Sciences Education*, 7: 85- 97.

Carrillo-de-la-Peña, M.T., Baillès, E., Caseras, X., Martínez, A., Ortet, G. & Pérez, J. (2009). Formative assessment and academic achievement in pre-graduate students of health sciences. *Advances in Health Sciences Education*, 14: 61- 67.

Coderre, S., Woloschuk, W. & McLaughlin, K. (2009). Twelve tips for blueprinting. *Medical Teacher*, 31: 359- 361.

Cook, D.A. & Beckman, T.J. (2009). Does scale length matter? A comparison of nine-versus five-point rating scales for the mini-CEX. *Advances in Health Sciences Education*, 14: 655- 664.

Cook, D.A., Brydges, R., Ginsburg, S. & Hatala, R. (2015). A contemporary approach to validity arguments: a practical guide to Kane's framework. *Medical Education* 49: 560– 575.

Crisp, G.T. (2012). Integrative assessment: reframing assessment practice for current and future learning. *Assessment & Evaluation in Higher Education*, 37 (1): 33- 43.

Dahllöf, G., Tsilingaridis, G., & Hindbeck., H. (2004). A logbook for continuous self-assessment during 1 year in paediatric dentistry. *European Journal of Paediatric Dentistry*, 3: 163- 169.

Downing, S.M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, 37: 830- 837.

Downing, S.M. (2004). Reliability: on the reproducibility of assessment data. *Medical Education*, 38: 1006- 1012.

Downing, S.M. (2005). Threats to the validity of clinical teaching assessments: what about rater error? *Medical Education*, 39: 353- 355.

Downing, S.M. & Yudkowsky, R. (2009). Introduction to assessment in Health Professions. *Assessment in Health Professions Education*, Routledge, New York, pp 1- 20.

Driessen, E., van der Vleuten, C., Schuwirth, L., van Tartwijk, J. & Vermunt, J. (2005). The use of qualitative research for portfolio assessment as an alternative to reliability evaluation: a case study. *Medical Education* 39: 214- 220.

Durning, S.J., Cation, L.J., Markert, R.J. & Pangaro, L.N. (2002). Assessing the reliability and validity of the Mini-Clinical Evaluation Exercise for internal medicine residency training. *Academic Medicine*, 77 (9): 900- 904.

Epstein, R.M. (2007). Assessment in medical education. *The New England Journal of Medicine*, 356 (4).

Epstein, R.M. & Hundert, E.M. (2002). Defining and assessing clinical competence. *JAMA* 287 (2): 226- 235.

Falchikov, N. (2003). Involving students in assessment. *Psychology Learning and Teaching*, 3 (2): 102- 108.

Frank, J.R., Snell, L.S., Cate, O.T., Holmboe, E.S., Carraccio, C., Swing, S.R., Harris, P., Glasgow, N.J., Campbell, C., Dath, D., Harden, R.M., Iobst, W., Long, D.W., Mungroo, R., Richardson, D.L., Sherbino, J., Silver, I., Taber, S., Talbot, M., & Harris, K.A. (2010). Competency-based medical education: theory to practice. *Medical Teacher*, 32 (8): 638- 645.

Gezie, A., Khaja, K., Chang, V.N., Adamek, M.E. & Johnsen, M.B. (2012). Rubrics as a tool for learning and assessment: What do baccalaureate students think? *Journal of Teaching in Social Work*, 32: 421- 437.

Gibson, K. & Shaw, C.M. (2010). Assessment of active learning. *The International Studies Encyclopaedia*. Denmark. Robert A. Blackwell Publishing, 2010. Available from:

ebs.wichita.edu/depttools/depttoolsmemberfiles/carolynshaw/Gibson%20Shaw%20compendium.pdf [Accessed 14 October 2017].

Govaerts, M., Schuwirth, L.W.T., van der Vleuten, C.P.M. & Muijtens, A.M.M. (2011). Workplace-based assessment: effects of rater expertise. *Advances in Health Sciences Education* 16: 151-165.

Govaerts, M. & van der Vleuten, C.P.M. (2013). Validity in work-based assessment: expanding our horizons. *Medical Education*, 47: 1164- 1174.

Govaerts, M.J.B., Van de Wiel, M.W.J., Schuwirth, L.W.T., van der Vleuten, C.P.M. & Muijtjens, A.M.M. (2013). Workplace-based assessment: raters' performance theories and constructs. *Advances in Health Sciences Education* 18: 375- 396.

Harris, P., Bhanji, F., Topps, M., Ross, S. Lieberman, S., Frank, J.R., Snell, L. & Sherbino, J. (2017). Evolving concepts of assessment in a competency-based world. *Medical Teacher*, 39 (6): 603- 608.

Hatala, R., Ainslie, M., Kassen, B.O., Mackie, I. & Roberts, M. (2006). Assessing the mini-Clinical Evaluation Exercise in comparison to a national specialty examination. *Medical Education*, 40: 950- 956.

Hatala, R. & Norman, G.R. (1999). In-training evaluation during an internal medicine clerkship. *Academic Medicine*, 74 (10): S118- S120.

Hays, R.B., Hamlin, G. & Crane, L. (2015). Twelve tips for increasing the defensibility of assessment decisions. *Medical Teacher*, 37: 433- 436.

Health Professions Council of South Africa (HPCSA). (2014). Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa. Available from: http://www.hpcsa.co.za/uploads/editor/UserFiles/downloads/medical_dental/MDB%20Core%20Competencies%20-%20ENGLISH%20-%20FINAL%202014.pdf [Accessed 21 July 2017].

Hernández, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education*, 64: 489- 502.

Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2: 130- 144.

Kane, M.T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement* 50 (1):1- 73.

Konopasek, L., Norcini, J. & Krupat, E. (2016). Focusing on the formative: Building an assessment system aimed at student growth and development. *Academic Medicine*, 91 (11): 1492-1497.

Krasne, S., Wimmers, P.F., Relan, A. & Drake, T.A. (2006). Differential effects of two types of formative assessment in predicting performance of first year medical students. *Advances in Health Sciences Education*, 11: 155- 171.

Kreiter, C.D. & Ferguson, K.J. (2001). Examining the generalizability of ratings across clerkships using a clinical evaluation form. *Evaluation & the Health Professions*, 24: 36-46.

Lara, R.F., Mogensen, K.M. & Markuns, J.F. (2016). Effective Feedback in the Education of Health Professionals. *Support Line*, 38 (2): 3-8.

Lau, A.M.S. (2016). 'Formative good, summative bad?'- A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, 40 (4): 509- 525.

McDade, S.A. (1995). Case study pedagogy to advance critical thinking. *Teaching of Psychology* 22 (1).

McManus, I.C., Elder, A.T. & Dacre, J. (2013). Investigating possible ethnicity and sex bias in clinical examiners: an analysis of data from the MRCP (UK) PACES and nPACES examinations. *BMC Medical Education* 13: 103.

Mertler, C.A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7 (25).

Miller, G.E. (1990). The assessment of clinical skills/ competence/ performance. *Academic Medicine* 65: S63- 67.

Morrison, S. & Free, K.W. (2001). Writing multiple choice test items that promote and measure critical thinking. *Journal of Nursing Education*, 40 (1): 17- 24.

Nicholson, P., Gillis, S. & Trisha Dunning, A.M. (2009). The use of scoring rubrics to determine clinical performance in the operating suite. *Nurse Education Today*, 29: 73-82.

Norcini, J.J., Blank, L.L., Duffy, F.D. & Fortna, G.S. (2003). The Mini-CEX: A method for assessing clinical skills. *Annals of Internal Medicine*, 138: 476- 481.

Norcini, J. & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher*, 29: 9-10, 855- 871.

Patil, S.Y., Gosavi, M., Bannur, H.B. & Ratnakar, A. (2015). Blueprinting in assessment: A tool to increase the validity of undergraduate written examinations in pathology. *International Journal of Applied Basic Medicine Research* 5 (Suppl1): S76- S79.

Paukert, J.L., Richards, M.L. & Olney, C. (2002). An *encounter card system for increasing feedback to students*. *The American Journal of Surgery*, 183 (3): 300- 304.

Pelgrim, E.A.M., Kramer, A.W.M., Mokkink, H.G.A., van den Elsen, L., Grol, R.P.T.M & van der Vleuten, C.P.M. (2011). In-training assessment during direct observation of single-patient encounters: a literature review. *Advances in Health Sciences Education*, 16: 131- 142.

Perlman, C.L. (2003). Performance assessment: designing appropriate performance tasks and scoring rubrics. Available from: <http://files.eric.ed.gov/fulltext/ED480070.pdf> [Accessed 14 October 2017].

Popil, I. (2011). Promotion of critical thinking by using case studies as teaching method. *Nurse Education Today* 31: 204- 207.

Riaz, F., Yasmin, S. & Yasmin, R. (2015). Introducing regular formative assessment to enhance learning among dental students at Islamic International Dental College. *JPMA*, 65: 1277- 1282.

Royal, K.D. (2017). Four tenets of modern validity theory for medical education assessment and evaluation. *Advances in Medical Education and Practice* 8: 567- 570.

Sadler, D.R. (1989).Formative assessment and design of instructional systems. *Instructional Science*, 18: 119-144.

South African Qualifications Authority (SAQA). (2001). National policy and criteria for designing and implementing assessment for NQF qualifications and part-qualifications and professional designations in South Africa. Available from:

<http://www.saqa.org.za/docs/pol/2015/National%20Policy%20for%20Assessment.pdf>
[Accessed 21 July 2017].

SAQA. (2005). Guidelines for integrated assessment. Available from:
<http://www.saqa.org.za/docs/guide/2005/intassessment.pdf> [Accessed 26 August 2017]

Schuwirth, L.W.T. & van der Vleuten, C.P.M. (2010). How to design a useful test: the principles of assessment. In: Stanwick, T (ed). *Understanding Medical Education: Evidence, Theory and Practice*. Wiley-Blackwell, John Wiley & Sons, Ltd.

Stupart, D., Goldberg, P., Krige, J. & Khan, D. (2008). Does examiner bias in undergraduate oral and clinical surgery examinations occur? *SAMJ*, 98 (10): 805- 807.

Tanner, C.A. (2006). Thinking like a nurse: A research-based model of clinical judgement in nursing. *Journal of Nursing Education*, 45 (6): 204- 211.

Taylor D.C.M & Hamdy, H. (2013) Adult learning theories: Implications for learning and teaching in medical education: AMEE guide No. 83. *Medical Teacher*, e1-e12.

The Royal College of Physicians and Surgeons of Canada. *CanMEDS Physician competency framework*. (2015). Available from:
http://canmeds.royalcollege.ca/uploads/en/framework/CanMEDS%202015%20Framework_EN_Reduced.pdf [Accessed 23 March 2017].

Tziner, A., Murphy, K.R. & Cleveland J.N. (2005). Contextual and rater factors affecting rating behaviour. *Group & Organization management*, 30 (1): 89- 98.

University of the Western Cape (UWC). (2017). *Assessment Policy*. General Calendar pp 50-53.

Unluer, S. (2012). Being an insider researcher while conducting case study research. *The Qualitative Report*, 17 (58): 1- 14.

van der Vleuten, C.P.M. (1996). The assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education*, 1: 41- 67.

van der Vleuten, C.P.M. & Schuwirth L.W.T. (2005). Assessing professional competence: From methods to Programmes. *Medical Education*, 39: 309- 317.

van der Vleuten, C.P.M., Schuwirth L.W.T., Driessen, E.W., Govaerts, M.J.B. & Heeneman, S. (2015). Twelve tips for programmatic assessment. *Medical Teacher*, 37: 641- 646.

Wass, V., van der Vleuten, C., Shatzer, J. & Jones, R. (2001). Assessment of clinical competence. *The Lancet*, 357: 945- 949.

Wimmers, P. (2006). Developing clinical competence. Printpartners Ipskamp, B.V., Netherlands. Available from:
https://www.researchgate.net/publication/254804646_Developing_Clinical_Competence [Accessed 2 October 2017].

Wood, D.F. (2010). Formative assessment. In: Stanwick, T (ed). *Understanding Medical Education: Evidence, Theory and Practice*. Wiley-Blackwell, John Wiley & Sons, Ltd.

Yorke, M. (2003). Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45: 477- 501.

APPENDIX A: Minimal procedural requirements/ quotas

Procedure	At end of 5th year (cumulative quota over 4th and 5th years)
Diagnosis & Treatment Planning	10
Prophy / Fluoride Treatment	10
Diet Analysis	3
Bleaching / Micro Abrasion	1
GA attendance	1
Emergency Treatment	5
Permanent Restorations of 1° teeth (including ART)	20
Pulp Treatments	2
Steel Crown / Strip Crowns	2
Fissure Sealants (1° & 2° teeth)	30
Permanent Restorations of 2° teeth	10
Partial Denture / Space Maintainers	1
Trauma	1
Completed Cases	6
Sedation / N ₂ O	1
Comprehensive case presented	1
Unknown case presented	1
Special needs case Hist, Diag, treatment plan	1

APPENDIX B: Module descriptor

Faculty	Dentistry
Module Topic	Paediatric Dentistry V
Alpha-numeric Code	PED 511 & 512
NQF Level	8
Duration	1 year
Programmes in which the module will be offered	BChD (5101)
Year level	5
Main Outcomes	<p>On completion of this module, students should be able to:</p> <ul style="list-style-type: none"> • Integrate the principles of behaviour management and apply them to the comprehensive management of the child. • Formulate a diagnosis and comprehensive treatment plan. • Manage paediatric patients under general anaesthesia and sedation. • Identify the need to refer children for specialist pharmacotherapeutic management. • Adapt a comprehensive preventive strategy to the child's individual needs. • Treat caries using relevant restorative techniques and materials based on the current evidence. • Recognize the indications and contraindications for all pulp therapy procedures in a child. • Perform pulp therapy procedures on primary and permanent teeth. • Identify malocclusions and recognize the need for interceptive treatment and/ or orthodontic referral. • Construct appliances for space maintenance and/or refer patients appropriately. • Distinguish between and manage different types of dental trauma. • Manage patients requiring more specialized care i.e. Patients with medical, mental and physical disabilities and/ or diseases. • Recognize and appropriately report suspected cases of child abuse and neglect. • Refer the paediatric patient to the appropriate health care provider when necessary.
Main Content	<ul style="list-style-type: none"> • Integrated patient case presentations • Unknown patient case • Patient management (including 50% quota)
Pre-requisite modules	4 th year content (see Study Guide)
Methods of Student Assessment	<p>Continuous Assessment (CA): 60% clinical + 40% theory</p> <p>Final Assessment (FA): 60% continuous + 40% exam</p>

Student name

Year:

Student code

Staff name

Staff code

Date:

Rating	%
1	20
2	40
3	50
4	60
5	75+

**Final mark
(rubric)**

**Pt
1**

**Pt
2**

Pt
3

Additional comments:

Procedures assisted

Clinical

Theory

Supervisor's signature: Student's signature:

Mark	Clinical (includes professionalism)	Theory
1	<p><i>Patient bookings/ professionalism</i></p> <ul style="list-style-type: none"> • Student absent/ Cancels patients without permission • No patient booked by student at least a week in advance <p><i>Clinical</i></p> <ul style="list-style-type: none"> • Detrimental to patient—mistake not rectifiable • Missed important information eg. History of pain • Missed obvious pathology • Treatment plan unacceptable • Wrong procedure performed • Starting procedure without permission or signed treatment plan • Restorations clinically unacceptable 	<ul style="list-style-type: none"> • Complete lack of core knowledge—could not answer any of the questions posed <ul style="list-style-type: none"> ➢ Not able to justify material selection despite prompting ➢ Cannot integrate theory with clinical practice ➢ Complete lack of critical thinking ➢ Cannot motivate treatment chosen
2	<p><i>Patient bookings/ professionalism</i></p> <ul style="list-style-type: none"> • Double-booking patients without permission • Arriving/ starting late • Lack of time management/ poor planning • Leaving without permission (if patient doesn't arrive) • Dismissing patient without supervisor's permission • Non-compliance with dress code <p><i>Clinical</i></p> <ul style="list-style-type: none"> • Instrument seal broken before patient arrives • Poor infection control, untidy cubicle, gloves not in biohazard container • Could not correlate history and clinical picture • Took radiographs before doing a clinical charting • Starting tooth preparation without radiographs where indicated • Missed no pathology clinically but treatment plan unacceptable • Quality of procedure unacceptable but mistake rectifiable • Restorations need major adjustment 	<ul style="list-style-type: none"> • Needs prompting to justify material selection/ procedures • Could answer less than half of questions posed <ul style="list-style-type: none"> ➢ Has some idea of rationale for treatment ➢ Some core knowledge ➢ Critical thinking skills/ reasoning not sound
3	<ul style="list-style-type: none"> • Incomplete History/ Diagnosis • OH and Fluoride only • Reasonable history taking-- missed some things but could <u>not</u> answer when prompted • Treatment plan acceptable but needs major revision • Clinical work acceptable but guidance required with procedure • Clinically acceptable restorations but needs minor adjustment 	<ul style="list-style-type: none"> • Core knowledge acceptable/average <ul style="list-style-type: none"> ➢ Reasonable insight regarding procedures/ materials ➢ Fair idea for choosing treatment option ➢ Theoretical knowledge and critical thinking skills sound
4	<ul style="list-style-type: none"> • Missed some things regarding history but could answer when prompted • Treatment plan acceptable. Needs only minor revision. • Very limited guidance needed with procedure • Good quality restoration (good contacts/ marginal adaptation)-- minor adjustment needed 	<ul style="list-style-type: none"> • Sound knowledge (better than average) <ul style="list-style-type: none"> ➢ Good motivation but doesn't cover all the possible treatment options ➢ Competent ➢ Able to justify material selection ➢ Good insight demonstrated but can improve ➢ Theoretical knowledge good
5	<ul style="list-style-type: none"> • Missed nothing. Covered all the bases without prompting • Diagnosis and treatment plan spot-on • No guidance needed with procedure • Exceptional patient management • Perfect restoration (good contacts/ marginal adaptation/ no overhangs)—no adjustment needed • All restorations polished—no rough spots 	<ul style="list-style-type: none"> • Substantial knowledge (more than expected) <ul style="list-style-type: none"> ➢ Good motivation— aware of all the treatment possibilities/ options ➢ Displays in-depth understanding ➢ Able to justify material selection/ choice of procedures critically ➢ Excellent insight demonstrated ➢ Theoretical knowledge excellent

NB! Lowest mark achieved gets recorded

APPENDIX D: Excerpt of detailed module descriptor used to compile clinical rubric

Module outcomes	<p>On completion of this module, the student must demonstrate competence in the following knowledge, skills, and values</p> <p><u>Knowledge-</u> Demonstrate an understanding of:</p> <ul style="list-style-type: none"> • The normal and abnormal physical and psychological development of the child. • When to refer/ treat patients for pharmacotherapeutic intervention (sedation/ GA) • the various stages of decay and how they present • the morphologic differences between primary and permanent teeth and the impact it has on restorative procedures • the indications and contraindications for all pulp therapy procedures in a child • malocclusions • the different types of dental trauma • normal and abnormal anatomy of the oral cavity <p><u>Skills-</u>The student must be able to:</p> <ul style="list-style-type: none"> • Formulate a diagnosis and comprehensive treatment plan taking the patient's treatment needs into consideration • Apply the principles of behaviour management • Select suitable preventive measures that are relevant to each clinical situation • Plan a preventive strategy tailored to the patient's needs • Recognize and manage cavitated lesions/ various states of dental decay using relevant restorative techniques • Select the appropriate restorative materials and motivate choice of materials • Perform pulp therapy procedures on primary and permanent teeth • Manage different types of dental trauma • Recognize the need for interceptive treatment and/ or orthodontic referral • Construct appliances for space maintenance • Provide comprehensive dental care for the more common oral and dental diseases in the child patient. • Management patients requiring more specialized care holistically i.e. patients with medical, mental and physical disabilities and/ or diseases • Recognize and appropriately report suspected cases of child abuse and neglect • Refer patients to the appropriate health care provider when necessary <p><u>Values-</u> The student must be able to:</p> <ul style="list-style-type: none"> • Communicate with the paediatric patient and the parent/ caregiver as well as other health care providers in a professional manner. • Promote an understanding of the cultural influences on feeding practices and dietary recommendations • Be sensitive to the impact of cultural and social aspects on patient management • Promote an understanding of the ethical implications of treating/ not treating oral disease • Manage cases of abuse and neglect confidentially and with the necessary sensitivity
------------------------	---

APPENDIX E: Analytic rubric (Blinded patient case)

CATEGORY	MARK OUT OF 10	DETAIL
History taking	4 and under	<ul style="list-style-type: none"> Missed important information eg. History of pain
	5	<ul style="list-style-type: none"> Could not correlate history and clinical picture
	6	<ul style="list-style-type: none"> Reasonable history taking. Missed some things but could not answer when prompted.
	7	<ul style="list-style-type: none"> Missed some things but could answer when prompted
	8+	<ul style="list-style-type: none"> Missed nothing. Covered all the bases without prompting
Diagnosis and treatment plan	4 and under	<ul style="list-style-type: none"> Missed obvious pathology. Treatment plan unacceptable
	5	<ul style="list-style-type: none"> Missed no pathology clinically. Treatment plan unacceptable
	6	<ul style="list-style-type: none"> Treatment plan acceptable. Needs major adjustment
	7	<ul style="list-style-type: none"> Treatment plan acceptable. Needs minor adjustment
	8+	<ul style="list-style-type: none"> Diagnosis and treatment plan spot-on
Rationale for treatment plan	4 and under	<ul style="list-style-type: none"> Cannot motivate treatment chosen
	5	<ul style="list-style-type: none"> Has some idea
	6	<ul style="list-style-type: none"> Fair idea for choosing treatment option
	7	<ul style="list-style-type: none"> Good motivation. Doesn't cover all the possible treatment options.
	8+	<ul style="list-style-type: none"> Good motivation Is aware of all the treatment possibilities/ options
Knowledge and insight	4 and under	<ul style="list-style-type: none"> Cannot integrate theory with clinical practice Theoretical knowledge poor Complete lack of critical thinking
	5	<ul style="list-style-type: none"> Theoretical knowledge fair Critical thinking skills fair
	6	<ul style="list-style-type: none"> Theoretical knowledge sound Critical thinking skills sound
	7	<ul style="list-style-type: none"> Theoretical knowledge good Critical thinking skills good
	8+	<ul style="list-style-type: none"> Theoretical knowledge excellent Critical thinking skills excellent

APPENDIX F: Ethics approval (US)

68

N. Mohamed 05-07-2017



Health Research Ethics Committee (HREC)

Approval Notice

New Application

06-Mar-2018

Ethics Reference #: X18/02/002

Title: Workplace-based assessment: Is it a true reflection of clinical competence of final year students in Paediatric Dentistry

Dear Dr N Mohamed

Your application for ethics exemption dated 13 February 2018 refers.

The Health Research Ethics Committee has assessed your application and considers this research proposal to meet the requirements for ethics exemption. This project involves a quality assurance audit using existing records.

This letter confirms that this research is now registered and you can proceed with study related activities.

Yours sincerely,

Franklin Weber

HREC Coordinator,

Health Research Ethics Committee 1

Health Research Ethics Council (NHREC) Registration Number:

REC-130408-012 (HREC1)•REC-230208-010 (HREC2)

Federal Wide Assurance Number: 00001372

*Office of Human Research Protections (OHRP) Institutional Review Board (IRB) Number:
IRB0005240 (HREC1)•IRB0005239 (HREC2)*

The Health Research Ethics Committee (HREC) complies with the SA National Health Act No. 61 of 2003 as it pertains to health research. The HREC abides by the ethical norms and principles for research, established by the [World Medical Association \(2013\). Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects](#); the South African Department of Health (2006). [Guidelines for Good Practice in the Conduct of Clinical Trials with Human Participants in South Africa \(2nd edition\)](#); as well as the Department of Health (2015). [Ethics in Health Research: Principles, Processes and Structures \(2nd edition\)](#).

The Health Research Ethics Committee reviews research involving human subjects conducted or supported by the Department of Health and Human Services, or other federal departments or agencies that apply the Federal Policy for the Protection of Human Subjects to such research (United States Code of Federal Regulations Title 45 Part 46); and/or clinical investigations regulated by the Food and Drug Administration (FDA) of the Department of Health and Human Services.

APPENDIX G: Ethics approval (UWC)



07 March 2018

RE: REQUEST FOR PERMISSION TO CONDUCT RESEARCH AT THE UNIVERSITY OF THE WESTERN CAPE

Name of Researcher	: Nadia Mohamed
Research Topic	: Workplace-based assessment: Is it a true reflection of clinical competence of final year students in Paediatric Dentistry
Date of issue	: 07/03/2018
Reference number	: UWCRP070318NM

This serves as acknowledgement that you have obtained and presented the necessary ethical clearance and your institutional permission required to proceed with the above referenced project.

Approval is granted for you to conduct research at the University of the Western Cape for the period 7 March 2018 to 28 February 2019. You are required to engage this office in advance if there is a need to continue with research outside of the stipulated period. The manner in which you conduct your research must be guided by the conditions set out in the annexed agreement: *Conditions to guide research conducted at the University of the Western Cape*.

The University of the Western Cape promotes the generation of new knowledge and supports new research. It also has a responsibility to be sensitive to the rights of the students and staff on campus. This office will require of you to respect the rights of students and staff who do not wish to participate in interviews and/or surveys.

It is also incumbent on you to first furnish this office with a copy of the proposed publication should you wish to reference the University's name, spaces, identity, etc. prior to public dissemination.

Please be at liberty to contact this office should you require any assistance to conduct your research or specifically require access to either staff or student contact information.

Yours sincerely

DR AHMED SHAIKJEE
DEPUTY REGISTRAR: ACADEMIC ADMINISTRATION
OFFICE OF THE REGISTRAR

UWCRP070318NM
Page 1 of 3