

**AN INVESTIGATION INTO THE VALIDITY AND RELIABILITY OF AN
INSTRUMENT FOR THE ASSESSMENT OF CLINICAL PERFORMANCE
DURING WORK INTEGRATED LEARNING OF EMERGENCY MEDICAL
CARE STUDENTS AT THE UNIVERSITY OF JOHANNESBURG.**

By

BERNARD VAN TONDER

**A research report submitted in fulfilment of the requirements for the
degree Master of Emergency Medical Care in the Faculty of Health
and Wellness Sciences at the Cape Peninsula University of
Technology**

Supervisors: Professor C. Vincent-Lambert, Mr. N. Naidoo

Cape Town

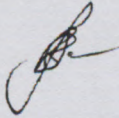
October 2016

CPUT copyright information

The dissertation/thesis may not be published either in part (in scholarly, scientific or technical journals), or as a whole (as a monograph), unless permission has been obtained from the University

DECLARATION

I, Hermanus Bernardus van Tonder, declare that this research report is my own work. It is being submitted for the degree of Master of Emergency Medical Care at the Cape Peninsula University of Technology. It has not been submitted before for any degree at this or any other University.



October 2016

ABSTRACT

Background - As emergency medical care students approach the exit level of their four-year qualification, additional focus get placed on assessment of their ability provide patient care in the real world pre-hospital emergency care environment. Upon graduation, there is no opportunity for newly graduated emergency care practitioners to complete an internship programme. The assessment of clinical competence is therefore regarded as a critically important and invaluable activity within the academic unit. Academic staff within the Emergency Medical Care department at UJ recognised the need for the development of a standardised assessment instrument to purposefully assess pre-hospital clinical performance and developed an assessment instrument referred to as the University of Johannesburg Clinical Performance Assessment Instrument (UJ CPAI). Having developed the UJ CPAI it became necessary and important to scientifically investigate and evaluate the extent to which the CPAI (as a newly developed instrument) meets the requirements of what is considered to be a "good assessment instrument". For this reason investigation of the validity, reliability and end-user support for the implementation of the UJ CPAI became the central aim and focus of this study.

Method – A prospective, descriptive, sequential exploratory mixed-method design was chosen for this study. The research methods included a literature review, administration of a survey questionnaire assessing the construct and content validity of the CPAI, analysis of quantitative data gathered from completed UJ CPAI mark sheets, analysis of qualitative data generated by a focus group discussion and method triangulation and corroboration of quantitative and qualitative data sets.

Results – This study highlighted some interesting similarities and differences in the assessment instruments used for the assessment of clinical performance. Central findings included that the construct validity of the UJ CPAI needs to be improved in certain areas, that the content validity is shown to be satisfactory and that the instrument was capable of producing results

showing good inter-rater reliability more than half the time. The study shows support from educators for the implementation of an instrument such as the UJ CPAI; however, in its current format the assessment process appears too cumbersome. Refinements and further study is needed to develop the instrument to a point where it can be considered scientifically validated and universally accepted as suitable for the assessment of clinical competence in the prehospital emergency care environments.

ACKNOWLEDGEMENTS

A project such as this cannot be completed alone. I take this opportunity to thank the following individuals who have been instrumental to the successful completion of this project.

I would like to thank my supervisor and mentor, Professor Craig Vincent-Lambert, for the invaluable comments, direction, and, sometimes stern, encouragement through the learning process of this project. Also a sincere word of thanks to my co-supervisor, Mr. N. Naidoo, for his input, advice and engagement in this project.

A special thanks to my family. Thank you to my wife, Tanya and my children, Armand and Christelle for all of the support and encouragement you offer to me.

To my colleagues from the Department Emergency Medical Care, thank you for your support and encouragement.

CONTENTS

DECLARATION.....	II
ABSTRACT	III
ACKNOWLEDGEMENTS.....	V
LIST OF FIGURES	XII
LIST OF TABLES.....	XIV
CHAPTER 1 –BACKGROUND AND CONTEXT	15
1.1 INTRODUCTION.....	15
1.2 BACKGROUND.....	15
1.3 CONCEPTUAL FRAMEWORK.....	17
1.4 BACKGROUND TO THE RESEARCH PROBLEM AND PROBLEM STATEMENT.....	19
1.5 RESEARCH QUESTIONS	20
1.6 AIM	20
1.7 STUDY OBJECTIVES.....	21
1.8 RESEARCH PARADIGM.....	21
1.9 RESEARCH DESIGN AND METHOD	22
1.10 RESEARCH METHODS	23
1.11 POSSIBLE BENEFITS.....	25
1.12 OUTLINE OF THE DISSERTATION AND CHAPTERS TO FOLLOW..	25
CHAPTER 2- LITERATURE REVIEW.....	28
2.1 INTRODUCTION.....	28
2.2 EMERGENCY CARE EDUCATION IN SOUTH AFRICA.....	28
2.3 TEACHING AND LEARNING IN HEALTH SCIENCES EDUCATION .	29
2.3.1 Meaningful learning	30
2.3.2 Clinical Learning in Health Sciences Education	34
2.3.3 Skills acquisition during clinical learning.....	35
2.4 ASSESSMENT IN HEALTH SCIENCES EDUCATION	36
2.4.1 Assumptions and Principles that underpin assessment.....	37
<i>Validity</i>	37
<i>Reliability</i>	38

<i>Credibility</i>	39
<i>Feasibility</i>	40
2.5 ASSESSMENT OF CLINICAL COMPETENCE.....	41
2.6 ASSESSMENT OF CLINICAL PERFORMANCE	45
2.6.1 Workplace-based assessment instruments used to assess clinical performance.....	47
2.7 CONCLUSION.....	49
CHAPTER 3 –RESEARCH METHODS AND PROCEDURES	51
3.1 INTRODUCTION.....	51
3.2 AIM.....	51
3.3 OBJECTIVES.....	51
3.4 RESEARCH PARADIGM.....	52
3.5 RESEARCH DESIGN.....	54
3.6 RESEARCH METHODS	56
Step 1- Literature Review.....	57
Step 2- Analysis and description of the Design, Format and Structure of the UJ CPAI.	57
Step 3 - Assessment of the validity of the CPAI in this research study.	59
Step 4 - Investigation of the Reliability of the CPAI by the Analysis of Results from Completed UJ CPAI Mark Sheets.	66
Step 5 - The focus group discussion.....	67
Steps 6 and 7 - Triangulation of data, Answer Research Question and Make Necessary Recommendations.	70
3.7 ETHICAL CONSIDERATIONS.....	71
3.8 VALIDITY, RELIABILITY AND TRUSTWORTHINESS OF THIS STUDY.....	72
3.9 SUMMARY.....	73
CHAPTER 4- CONSTRUCT VALIDITY, CONTENT VALIDITY AND INTER- RATER RELIABILITY	75
INTRODUCTION	75
4.1. PART ONE: FORMAT AND STRUCTURE OF THE UJ CPAI.....	76
4.1 .1 Determine the areas that need to be assessed.....	77

4.1.1 .1	Comparability of the assessment areas of the CPAI and other assessment instruments.	78
4.1.1 .2	Comparability of the UJ CPAI assessment areas to the components of a pre-hospital incident.	80
4.1.1 .3	Literature that validates the inclusion of the assessment areas in the CPAI	81
	Assessment area one - Preparation	82
	Assessment area two - Professional presentation	82
	Assessment area three - History taking, communication and general patient interaction	83
	Assessment area four - Clinical assessment and diagnostic skills	83
	Assessment area five - Decision making regarding clinical interventions	84
	Assessment area six - Clinical procedures, skills and treatment interventions	84
	Assessment area seven - Clinical documentation and record keeping	85
	Assessment area eight - Handover	85
	Assessment area nine - Ability to reflect on own practice, self-critique and accept clinical direction	86
4.1 .2	Determine Specific "Behaviours" to be evaluated for Each Area.	87
4.1 .3	Determining the Appropriate Answer Scale.	87
4.1 .4	Summary of Part One.	90
4.2.	PART TWO: CONSTRUCT VALIDITY OF THE UJ CPAI	91
4.2.1	Overall ratings awarded by all the participants in the survey questionnaire.	91
4.2.2	An Analysis of the Consistency of the Results.	92
4.2.3	General Comments on the Free Writing Component of the Questionnaire	94
4.2.4	Summary of Part Two	96
4.3	PART THREE: CONTENT VALIDITY OF THE UJ CPAI	97
4.3.1.	Summary of the Ratings Awarded by the participants	97
4.3.2.	A Comparison of the Ratings Awarded for Each Assessment Area	98

4.3.3. Summary of Part Three	99
4.4 PART FOUR: INTER- RATER RELIABILITY OF THE UJ CPAI	99
4.4.1 Results from the UJ CPAI mark sheets	99
4.4.2 The Frequency of Cronbach’s Alpha of More than 0.6	100
4.4.3 Inter-rater Reliability per Assessment Area	103
4.4.4 Summary of Part Four	107
CHAPTER 5- FOCUS GROUP DISCUSSION	111
5.1 INTRODUCTION.....	111
5.2 THE FOCUS GROUP DISCUSSION	111
5.4 PARTICIPANTS IN THE FOCUS GROUP DISCUSSION	112
5.5 THE FOCUS OF DISCUSSION FOR THE GROUP	113
5.6 THEMES AND TRENDS THAT EMERGED FROM THE FOCUS GROUP DISCUSSION	114
FOCUS AREA ONE- CONSTRUCT VALIDITY OF THE UJ CPAI.....	114
FOCUS AREA TWO- CONTENT VALIDITY OF THE UJ CPAI.....	120
FOCUS AREA THREE- INTER-RATER RELIABILITY OF THE UJ CPAI.....	123
FOCUS AREA FOUR- SUPPORT FOR THE FUTURE USE OF AN ASSESSMENT INSTRUMENT SUCH AS THE UJ CPAI.....	127
5.7 CONCLUSION.....	129
5.8 SUMMARY.....	130
CHAPTER 6 –DISCUSSION OF CENTRAL FINDINGS	132
6.1 INTRODUCTION.....	132
6.2 VALIDITY OF THE UJ CPAI.....	133
6.2.1 Construct validity of the UJ CPAI	134
i. Assessment area one – Preparation.....	136
ii. Assessment area two – Professional presentation.....	137
iii. Assessment area three – History taking, communication and general patient interaction.....	138
iv. Rating scale of the UJ CPAI.....	139
6.2.2 Content validity of the UJ CPAI.....	141
6.3 RELIABILITY.....	142

6.3.1	Construct validity	144
6.3.2	Inadequate Assessor training.....	144
6.3.3	Assessor bias.....	145
6.4	SUPPORT FOR THE IMPLEMENTATION OF THE UJ CPAI AS AN ASSESSMENT INSTRUMENT	146
6.4.1	Credibility.....	146
6.4.2	Feasibility	148
6.5	RECOMMENDATIONS REGARDING THE FUTURE USE OF THE CPAI WITHIN THE EMERGENCY MEDICAL CARE SETTING.	148
6.6	SUMMARY	149
	CHAPTER 7 – SUMMARY AND CONCLUSION	151
7.1	INTRODUCTION	151
7.2	SUMMARY OF THE RESEARCH STUDY AND DISSERTATION.....	151
7.3	RECOMMENDATIONS	153
7.4	POTENTIAL CONTRIBUTION OF THE STUDY	155
7.5	LIMITATIONS.....	155
7.6	PROPOSALS FOR FURTHER STUDIES	156
7.7	CONCLUSION	157
	REFERENCES	159
	ANNEXURE 1 – INVITATION TO PARTICIPATE IN THE STUDY.....	169
	ANNEXURE 2 – INFORMATION SHEET ON HOW TO COMPLETE THE UJ CPAI.....	171
	ANNEXURE 3 – UJ CPAI NARRATIVE	173
	ANNEXURE 4 – UJ CPAI MARK SHEET	192
	ANNEXURE 5– INVITATION TO COMPLETE A QUESTIONNAIRE.....	200
	ANNEXURE 6 – SELF-DESIGNED, NON-VALIDATED, EVIDENCE INFORMED QUESTIONNAIRE.....	202
	ANNEXURE 7 – INVITATION TO PARTICIPATE IN A FOCUS GROUP DISCUSSION	206
	ANNEXURE 8 – CONSENT FORM TO PARTICIPATE IN A FOCUS GROUP DISCUSSION	208
	ANNEXURE 9 – CONSENT FORM TO BE AUDIO RECORDED	209

ANNEXURE 10 – PLAN FOR THE FOCUS GROUP DISCUSSION	210
NNEXURE 11 – ONE-WAY ANOVA AS WELL AS TWO WAY RANDOM CHRONBACH'S ALPHA.....	213

LIST OF FIGURES

Figure 1 - Conceptual framework for this study	18
Figure 2 - Summary of research methods used in this study.....	24
Figure 3- Components of the Teaching, Learning and Assessment continuum.....	30
Figure 4- Fink's Taxonomy of Meaningful Learning. From Fink (2003). ...	34
Figure 5- Miller's Prism of Clinical Competence (Miller, 1990).	43
Figure 6- Research Methods used for this Study	56
Figure 7- Eden and Swoboda's Six Step Approach to the Development of Workplace-Based Assessment Tools. (Adapted from Edens and Swoboda, 2013).	77
Figure 8- Comparability of the assessment areas included in the UJ CPAI, the Mini-CEX, UK WPBA and Tavares' WBA instrument.....	80
Figure 9- Deconstruction of a routine pre-hospital incident.....	81
Figure 10- Marking scale of the UJ CPAI.....	89
Figure 11-Marking scale of the Mini-CEX	89
Figure 12- Marking scale of the UK WPBA	89
Figure 13- Marking scale of Tavares' WPBA.....	90
Figure 14- Comparison of Assessment Areas According to the Number of Good Ratings	99
Figure 15- Number of times a Cronbach's Alpha of 0.60 or higher was found, compared to the number of times a smaller coefficient was found.	101
Figure 16- Graphical representation of each Correlation Coefficient group	102
Figure 17- A representation of the Cronbach's Alpha for each Case Completed	103
Figure 18- Correlation Coefficient for each Assessment Area compared to the percentage of times the same mark was awarded.	106

Figure 19- A comparison between one-way ANOVA and two-way random
Methods when Calculating Cronbach's Alpha.107

LIST OF TABLES

Table 1- Likert-type scale ratings and their descriptors	62
Table 2- Ratings awarded by all seven participants for all nine assessment areas	91
Table 3 - Ratings awarded to Assessment area 8 by participants	92
Table 4- Ratings awarded by the participants to assessment areas two, three, four and six.	93
Table 5- A Comparison of Ratings Awarded by participants four and seven. A = Assessment Area.....	93
Table 6- Ratings awarded by all seven assessors for all nine assessment areas.	97
Table 7- Number of Good Ratings Awarded per Assessment Area	98
Table 8- The Number of Cases Where Assessors Agreed	104
Table 9- A Hypothetical Scenario Demonstrating the Value of Using r in Calculating Inter-Rater Reliability per Assessment Area.	105

CHAPTER 1 - BACKGROUND AND CONTEXT

1.1 INTRODUCTION

This study focused on an investigation and description of the validity and reliability of an instrument used in the assessment of clinical performance of emergency medical care students from the University of Johannesburg. This first chapter begins by introducing the background, context and conceptual framework underpinning the study. Thereafter, the problem statement, research question and central aim and objectives of the study are presented. Also included in Chapter 1 is a brief introduction to the research design and methodology, both of which are described in greater detail in Chapter 3. Chapter one concludes by providing a summary of the layout of the dissertation and the chapters that follow.

1.2 BACKGROUND

At the time of writing, the researcher was an academic staff member at the University of Johannesburg (UJ) Department of Emergency Medical Care (Department EMC). As part of his academic duties the researcher was responsible for activities associated with teaching, learning and assessment of third and fourth year EMC students during supervised clinical practice in the pre-hospital environment. It was during the performance of these duties where the researcher became increasingly interested in the area of clinical performance assessment.

As emergency medical care students approach the exit level of their four-year qualification, additional focus get placed on assessment of their ability to meaningfully apply previously mastered theoretical understandings and practical competencies into patient care in the real world pre-hospital emergency care environment. Upon graduation, there is

no opportunity for newly graduated emergency care practitioners to complete an internship programme. This means that unlike medical school graduates, EMC students will not have the opportunity to “find their feet” during mentored post qualification internship.

This means that Universities must be responsible to ensure that the students they graduate from their EMC programmes are clinically competent and ready for independent practice upon graduation. The assessment of clinical competence is therefore regarded as a critically important and invaluable activity within the academic unit. That said, there is currently no formal assessment of clinical performance in the pre-hospital environment, which is an integral and irreplaceable component of clinical competence (Miller, 2010).

Academic staff within the EMC department at UJ recognised the need for the development of a standardised assessment instrument to purposefully assess clinical performance. The assessment instrument would be able to guide the historically subjective approach to the assessment and judgment of clinical performance and in so doing produce more objective, fair, unbiased and valid results. During 2012 work on a standardised Clinical Performance Assessment Instrument (CPAI) [Annexure 3] was completed.

The UJ CPAI focused on the assessment of the following nine key areas:

- i) Preparation
- ii) Professional presentation
- iii) History taking, communication and general patient interaction
- iv) Clinical assessment and diagnostic skills
- v) Decision making regarding clinical interventions
- vi) Clinical procedures, skills and treatment interventions
- vii) Clinical documentation and record keeping

viii) Handover

ix) Ability to reflect on own practice, self-critique and accept clinical direction.

Clinical assessors observe the student as they management a patient in the prehospital environment and thereafter rate their performance in each of the nine areas using a five-point rating scale. Assessors award marks from 1 to 5 with 1 being poor and 5 the best possible performance. Assessor's allocations are guided by narrative that describes the typical actions that link best to each of the possible scores. The CPAI, its construction and application form the basis for this research and are discussed in greater depth in the chapters that follow.

Having developed the UJ CPAI it became necessary and important to scientifically investigate and evaluate the extent to which the CPAI (as a newly developed instrument) meets the requirements of what is considered to be a "good assessment instrument". In this regard Bashook indicates that assessments that assess clinical performance should be valid, reliable, feasible and credible (Bashook, 2005). Simply put, the implementation and continued use of the a newly developed assessment instrument such as the UJ CPAI will be dependent on it being shown to be a valid, reliable, feasible and credible tool. (Mitchell et al., 2011; Bashook, 2005). For this reason investigation of the validity, reliability and end-user support for the implementation of the UJ CPAI became the central aim and focus of this study.

1.3 CONCEPTUAL FRAMEWORK

A conceptual framework is regarded as a set of beliefs, theories and assumptions that guides the research process (Miles & Hubermann, 1994). Developing a conceptual framework is considered important because it

informs the study design and so assists the researcher to assess and refine the goals of the study, develop realistic and relevant research questions and select appropriate methods of study (Maxwell, 2005). The conceptual framework applicable to this study is presented in figure 1.1 below.

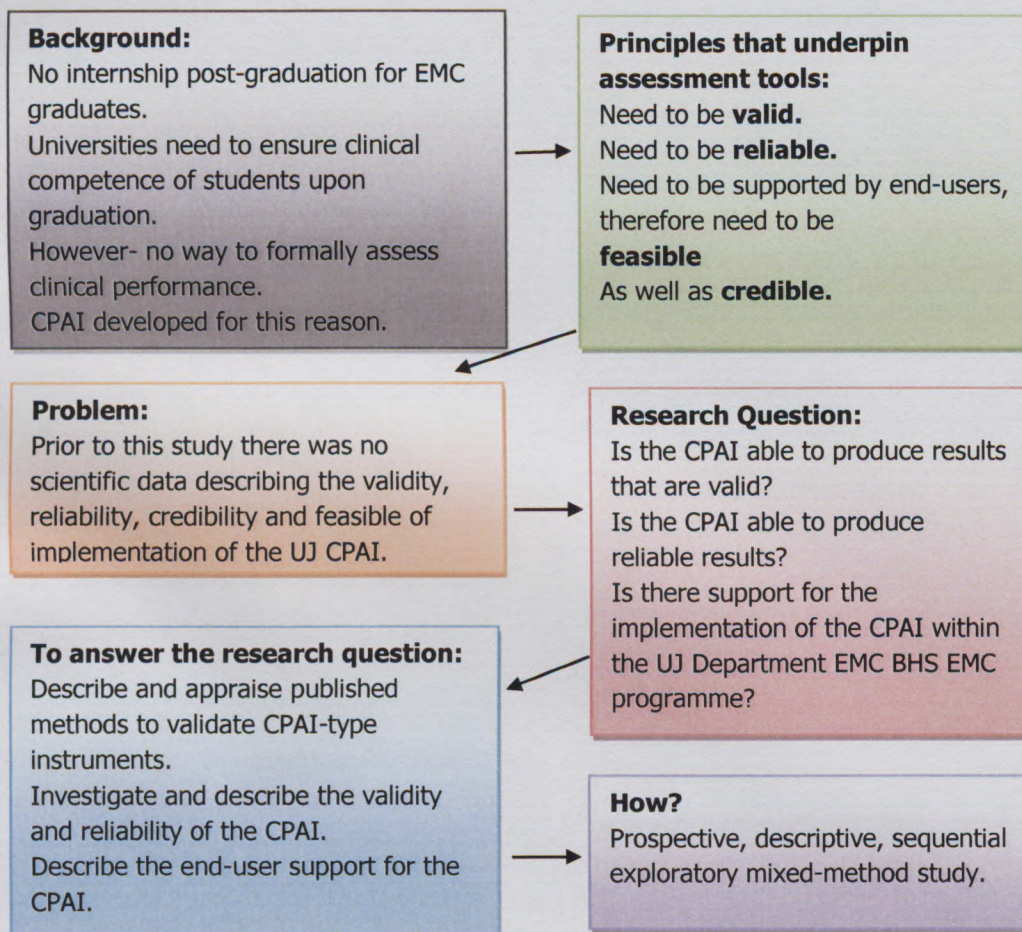


Figure 1 - Conceptual framework for this study

1.4 BACKGROUND TO THE RESEARCH PROBLEM AND PROBLEM STATEMENT.

As mentioned above, the CPAI (Annexure 3) was purposefully developed to assist in the assessment of clinical performance of Emergency Medical Care students during clinical learning. The problem faced was that prior to this study there was no empirical data describing:

a) The extent to which the newly developed CPAI is capable of producing valid and reliable results. The reason this is important is because academics need to ensure that the results produced by assessment instruments bear direct relevance to the outcomes being assessed. In the context of this study, the mark produced by the CPAI must give an accurate indication of the clinical performance of the student. Should the UJ CPAI not be a valid instrument, the results generated by its use will not be an accurate reflection of clinical performance and will therefore be of little value.

In addition to the above, the instrument will be used by numerous assessors. Investigation of the reliability of the instrument is therefore important as a good assessment instrument should lead to the same finding when used by more than one assessor to assess the same student.

b) The views and opinions of educators in the emergency care programme at UJ regarding the value and continued use of the CPAI. This refers to the feasibility and credibility of the CPAI. Instruments that are neither feasible nor credible will find little favour with assessors that need to use the instrument. Feasibility refers to matters such as simplicity and transparency. Credibility refers to matters such as procedural fairness and elimination of assessor bias (Bashook, 2005; Prober, 2010).

1.5 RESEARCH QUESTIONS

In light of the problem described above the central research questions became:

- a) What is the construct and content validity of the CPAI?
- b) To what extent is the CPAI capable of providing evidence that supports good inter-rater reliability?
- c) To what extent do staff members of the Department of Emergency Medical see the newly developed CPAI as capable of producing objective, valid, fair and reliable results?
- d) Is there evidence of a need for adjustments and or refinements to the CPAI and if so what are these?
- e) Finally (based on the answers to the above questions) is there sufficient scientific evidence to support the implementation and continued use of the CPAI within the UJ BHS EMC programme?

1.6 AIM

To address the research problem and answer the research questions described the following aim was derived:

To investigate and describe the validity and reliability of the CPAI as an instrument for the assessment of clinical performance of emergency medical care students in the pre-hospital environment.

1.7 STUDY OBJECTIVES

In pursuit of the aim described above the following research objectives were determined:

- a) To investigate and describe (within the context of available data and existing literature) the structure and format of the CPAI
- b) To investigate and describe methods used to validate clinical performance assessment instruments and identify appropriate methods to validate the CPAI in this study.
- c) To investigate and describe the construct and content validity of the CPAI.
- d) To calculate and report on the Inter-rater reliability of assessment results using the CPAI
- e) Investigate and describe the views and opinions of educators regarding the CPAI and its use for the the assessment of clinical performance of emergency medical care students in the pre-hospital environment and;
- f) To make recommendations regarding the structure, process, outcomes and future considerations for assessment of clinical practice within the UJ BHS EMC programme.

1.8 RESEARCH PARADIGM

Research paradigm may be compared to the lens through which the researcher views reality. It provides an organised method of thinking,

observing and interpreting findings within the setting of a specific discipline (Brink et al, 2012).

The researcher adopted a pragmatic approach in order to investigate and answer the research questions. Pragmatism draws on employing *what works* when answering the research question. The pragmatic theory accepts that all human activity stems from the need to solve problems, and places emphasis in employing methods that are effective in order to best solve or answer the research problem (Andrew & Halcomb, 2009) Through the lens of pragmatism, the researcher is able to value both objective as well as subjective knowledge (Morgan, 2007), which means that this paradigm is coherent with mixed method design described below (Morgan, 2007).

1.9 RESEARCH DESIGN AND METHOD

Research design is considered to be the logical action plan, or framework, which the researcher follows and uses to adequately connect gathered data and evidence in order to adequately address the research question (Yin, 1989; Flemke, 2004). A prospective, descriptive, sequential exploratory mixed-method design was chosen for this study. Mixed method designs may be defined as the use of quantitative research to assess the magnitude and/ or frequency of constructs, and then employing qualitative research to explore the meaning and understanding of constructs using more than one method and deliberately integrate these methods in order to draw on the strengths of each (Morgan, 2007; Andrew & Halcom; 2009). A mixed-method approach was seen to be suitable for this study, as this design allowed the researcher to value both qualitative as well as quantitative methods when attempting to answer the research question (Morgan, 2007; Hesse-Biber, 2010; Creswell, 2009).

The study had to be prospective in nature as data did not exist in this area prior to this study (Brink et al, 2012).

1.10 RESEARCH METHODS

As mentioned above, a sequential implementation sequence was followed during this study. The research steps and methods are described below. Firstly, the researcher conducted a literature review to describe the background to, and purpose for, the development of the CPAI within the context of available data and existing literature. The outcomes of this process are presented in the following chapter.

Following this, a self-designed survey questionnaire, that contained mainly closed questions, was circulated to all of the clinical assessors that used the UJ Clinical Performance Assessment Instrument (UJ CPAI) to assess students and agreed to participate in the study. This self-designed, non-validated, evidence-informed questionnaire was designed based on recommendations for testing validity content, published by Sireci and Faulkner – Bond (2014). The questionnaire explored the construct, as well as content validity of the CPAI. The content and construct validity was tested by calculating Cronbach's Alpha, using a one-way ANOVA as well as a two-way random model respectively. The questionnaire as well as the analysis of the responses will be discussed in greater detail in Chapters 3 and 4.

The next research step was a quantitative data analysis of assessment results captured by participants (clinical practice assessors) using the CPAI was conducted to calculate the inter-rater reliability of the CPAI using Cronbach's Alpha. This will also be elaborated upon in Chapter 4.

After the quantitative data had been analysed, views and opinions of educators regarding the CPAI and its use for the assessment of clinical performance of emergency medical care students in the pre-hospital environment was investigated. This fifth research step was done by a focus group discussion that generated qualitative data.

The data sets were analysed separately and confidence in the results were established. Findings from the above processes were then integrated using confirmatory triangulation in order to make recommendations regarding the structure, process, outcomes and future use of the CPAI within the emergency medical care setting. Figure 2 summarises the research methods used in this study.

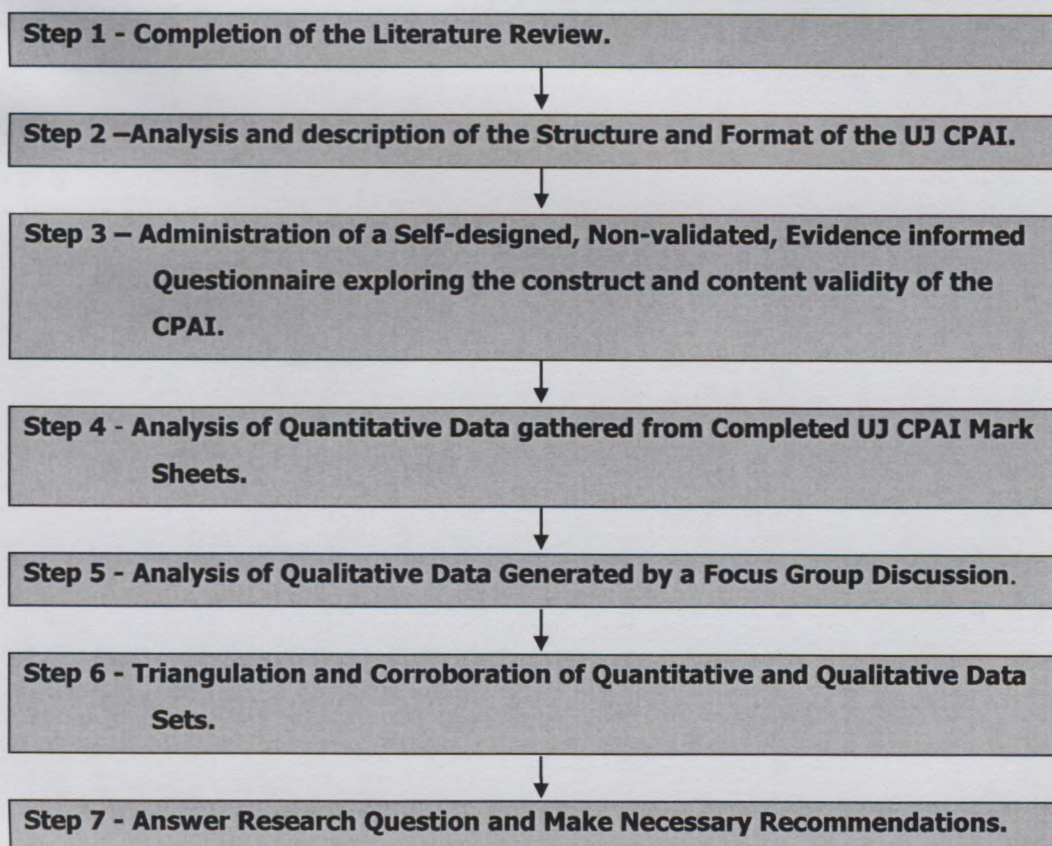


Figure 2 - Summary of research methods used in this study

Chapter 3 will provide more detail on the design and methods introduced above, defending their selection, application and importance in the context of this study.

1.11 POSSIBLE BENEFITS

The results and findings of this study have led to recommendations regarding future use and adjustments to the CPAI. In conducting this study the researcher provides new literature and insights into the assessment of clinical competence at UJ and in the local emergency medical care environment. Furthermore, this study provides a description of the considerations in assessing validity and reliability of tools that seek to assess and record clinical performance. Through publications and conference presentations existing knowledge in the area of pre-hospital clinical assessment of Emergency Medical Care students may be expanded. This in turn may lead to further debate and research into the implementation and utilization of similar instruments outside of the UJ and in the broader emergency care fields.

1.12 OUTLINE OF THE DISSERTATION AND CHAPTERS TO FOLLOW

This first Chapter provides the background and context to the study.

Chapter 2 presents literature on assessment practices in the pre-hospital emergency care education setting, as well as an exposition of the concept of validity and reliability determination.

Chapter 3 describes and defends the research design and methods that were applied in order to gather, analyse and interpret data related to the

aims and objectives of the study. The Chapter concludes by discussing ethical considerations applicable to this study.

Chapter 4 focuses on discussing the development of the clinical performance assessment instrument. In this chapter, the rationale behind the nine assessment outcomes the five-level rating scale of performance for each assessment area is explained. Chapter 4 also deals with the analysis of quantitative data sets derived from completed UJ CPAI mark sheets.

Chapter 5 deals with the focus group discussion used to investigate and describe the opinions of educators regarding the CPAI and its use for the assessment of clinical performance of emergency medical care students in the pre-hospital.

Chapter 6 focuses on corroborating the data sets and discussing the data in an attempt to answer the research question.

Chapter 7 concludes and summarises the study and its limitations before providing some recommendations for further study and future practice.

1.13 SUMMARY

This Chapter provided the background and context, aims, objectives of this study which was to investigate the validity and reliability of the CPAI as a newly developed clinical performance assessment instrument. The research problem, aim and objectives together with the mixed-methods sequential design and methodology were introduced and will be dealt with again in more detail in Chapter 3.

The following Chapter focuses on presenting literature on assessment practices in the pre-hospital emergency care education setting, as well as the concept of determining validity and reliability.

CHAPTER 2- LITERATURE REVIEW

2.1 INTRODUCTION

As mentioned in Chapter 1, the researcher aimed to provide a detailed critical description of the background to, and construction of, the UJ CPAI before investigating validity, reliability and end-user support for its use as a tool to assess clinical performance of Emergency Medical Care (EMC) students in the pre-hospital setting.

The researcher uses this chapter to present literature on emergency care education in South Africa and the importance of clinical teaching, learning and assessment within the pre-hospital emergency care environment. Theories that guide assessment, especially the assessment of clinical performance, are discussed including an argument for the need to validate an instrument for the assessment of clinical performance in the South African pre-hospital emergency care education domain. This literature review is important as it attempts to frame the study within the context of existing knowledge.

2.2 EMERGENCY CARE EDUCATION IN SOUTH AFRICA

The Emergency Care Practitioner credential is a 4-year, 480 credit, National Qualifications Framework (NQF) 8 Bachelor's Degree in Emergency Medical Care (BEMC). Graduates from the BEMC degree register with the Professional Board of Emergency Care (PBEC) at the Health Professions Council of South Africa (HPCSA) as Emergency Care Practitioners in independent practice (Vincent-Lambert, 2015). The idea is that assessment, clinical decision making, consultation and the development of Standard Operating Procedures and Protocols should in

the future to a large extent become owned and driven largely by ECPs (Vincent-Lambert, 2015).

Despite the high level of autonomous clinical responsibility expected from South African ECP graduates, there is currently no requirement for internship or community service prior to registration in the category of independent practice. This absence of post-graduation internship means that, unlike medical school graduates, graduates from B EMC programmes do not have the opportunity to "find their feet" during mentored post qualification clinical supervision. This abrupt transition from student to unsupervised, independent practice rightfully focuses serious attention on the teaching, learning and assessment of clinical competence of students prior to graduation.

2.3 TEACHING AND LEARNING IN HEALTH SCIENCES EDUCATION

In this section, the researcher will discuss some teaching and learning philosophies. Many of the modern philosophies are termed 'philosophies of learning'. Examples of such philosophies include Bloom's taxonomy of learning and Fink's taxonomy of meaningful learning. Although the taxonomy names or titles do not include the terms 'teaching' or 'assessment', the philosophies emphasise teaching and assessment as necessities to promote learning. The combination of teaching, learning and assessment is referred to by some as the Teaching, Learning and Assessment (TLA) continuum (Boud & Falchikov, 2007; Pintrige, 2002). Figure 3 represents the researcher's appreciation of the educational continuum as explained by Bloom (1956), Boud and Falchikov (2007), Pintrige (2002) and Kraftwohl (2002).

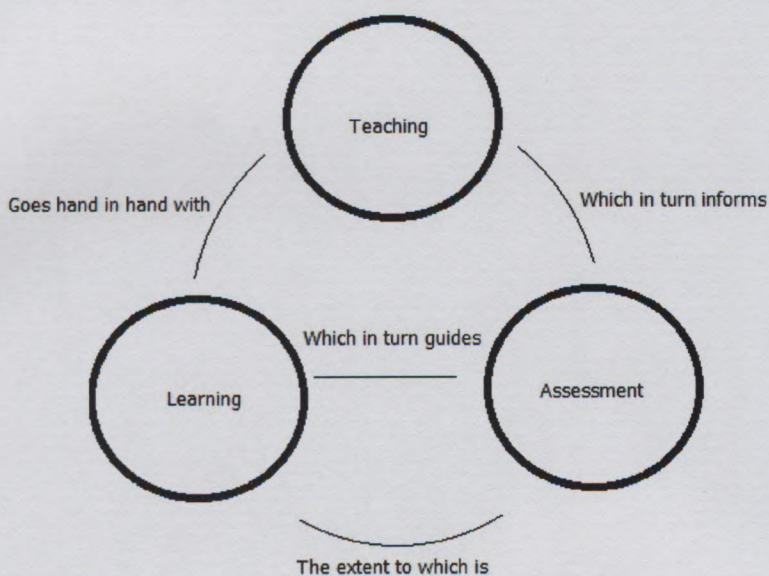


Figure 3- Components of the Teaching, Learning and Assessment continuum.

2.3.1 Meaningful learning

Teaching implies activity, not merely watching students interact with patients. The teacher therefore actively guides the students acting as an expert, role model and facilitator during clinical learning opportunities (Gaberson, Oemerann and Shellenbarger, 2015; Pratt and Magill, 1983). It is argued that the type of learning that occurs within the student cohort depends greatly on the didactic approaches and epistemological philosophies that underpin and inform the selection of a particular approach to teaching within a learning programme (Bloom, 1956; Isaacs, 1996).

In healthcare, the term 'learning' is frequently used to define a process where experiences, acquisition of knowledge and abilities generated long term behavioural change. De Sanctis defines learning as a process whereby individuals assimilate and gradually elaborate on increasingly

complex and abstract units of knowledge, such as concepts, categories, behavioural schemata or models and/or whereby behavioural skills and competences are mastered (De Sanctis, 2006; Cadorin et al., 2014). Unlike scientific theories, educational theories do not slot into clear-cut definitions. This is because these theories are usually the result of observation and assessment of teaching and learning practice based on an inductive approach (Dennick, 2012). An in-depth exploration of the different learning paradigms, philosophies and models is beyond the scope of this research study. The researcher will therefore limit this review of literature to philosophies applicable to the pre-hospital emergency care education context.

There is agreement amongst authors that teaching usually translate to learning. That said, however, authors also agree that the degree or extent to which learning takes place is not always sufficient. The extent to which learning takes place should be maximised by effective teaching strategies. For this reason, educators within the healthcare domain are currently rethinking traditional teacher-centred teaching and learning and are developing new ideologies and philosophies to move towards a student-centered approach (Stanley & Dougherty, 2010).

One such ideology of learning is Bloom's taxonomy, as revised by Kraftwohl in 2002. Bloom's original taxonomy, published in 1956, consisted of three domains, namely the factual, conceptual and procedural domains (Bloom, 1956; Kraftwohl, 2002; Pintrich, 2002). The criticism of Bloom's original taxonomy is that it follows a sequential model and is therefore not well suited for meaningful learning in modern healthcare education. Authors like Kraftwohl believe that meaningful learning is multidimensional in that it does not follow a sequential model. This means that relational and interactive and do not follow a hierarchical order.

Also, modern authors argue that another domain, that of metacognition, should be added (Kraftwohl, 2002; Pintrich, 2002). Metacognition, in summary, refers to the student's ability to take more control over their learning by knowing how they learn best and how they prepare best for exams. A simple example of this is to allow students access to outcomes and possible methods of assessment prior to an assessment taking place. Students are then able to gauge how prepared they are for the assessment by attempting to answer or meet the outcomes given to them. They can then either revisit content if they so wish or engage with new content (Kraftwohl, 2002; Pintrich, 2002; Mayer, 2002).

What makes the taxonomy as revised by Kraftwohl applicable to this research is that the argument is already made that students will learn more meaningfully if they know how they will be assessed, and which outcomes they will be assessed against. This also already grounds the argument that assessment forms part of a TLA continuum and is not something that happens only at the end of the year or module.

There are a number of other authors that also attempted to describe meaningful learning (Cadorin, Bagnasco, Rocco & Sasso, 2014; Johansen et al., 2012; Dreyfuerst, 2012; Fink, 2007). Although authors propose a number of different views and definitions, the central theme of meaningful learning is that new knowledge is constantly being evaluated and interpreted against existing knowledge, thereby creating more in-depth understanding (Johansen et al., 2012). In essence, this means that students learn best when constant feedback is given. In the case of this research study, feedback will be given with the use of the CPAI.

Also relevant to health care education, Dreyfuerst argues that meaningful learning goes deeper than memorizing information by promoting conceptual understanding and supporting the development of clinical

reasoning skills as a guide to professional practice (Dreyfuerst, 2012). An example of this statement is a concept called 'Debriefing in Meaningful Learning' (DML). Dreyfuerst argues that debriefing during clinical learning allows students to reflect on their actions. This reflection is guided by teachers and mentors and encourages students to apply thought processes and knowledge to clinical experience. This enhances their knowledge, decision-making skills and learning (Dreyfuerst, 2012). Coherence and cohesion between students' positive perception of the learning environment and practical proof of acquired learning are key to best teaching practices and stances of meaningful learning experience in students (Dreyfuerst 2012). Again, the envisaged role of the CPAI will be to aid the teacher in the debriefing and reflective process.

Another well-referenced taxonomy of learning is that of Fink (2007). Fink has developed a taxonomy that outlines six types of learning. These are basic knowledge, implementation, integration, human dimension, caring and learning to learn (Cadorin, Bagnasco, Rocco & Sasso, 2013). Unlike other approaches, such as Bloom's taxonomy for cognitive learning, the components of Fink's taxonomy are relational and interactive and do not follow a hierarchical order. Figure 4 illustrates Fink's taxonomy of meaningful learning.

As mentioned, the central theme of meaningful learning is that new knowledge is constantly being evaluated and interpreted against existing knowledge, thereby creating more in-depth understanding (Johansen et al., 2012). Cadorin, Bagnasco, Rocco & Sasso propose that, in addition to the central theme, meaningful learning also has three characteristics. These are i) Meaningful learning as an active construction process, ii) Meaningful learning as a change and iii) Meaningful learning developed through experience. In keeping with the context of this research study, which deals with clinical performance assessment, the researcher will

explore the third characteristic (meaningful learning through experience) within the context of the B.EMC Programme at the University of Johannesburg.

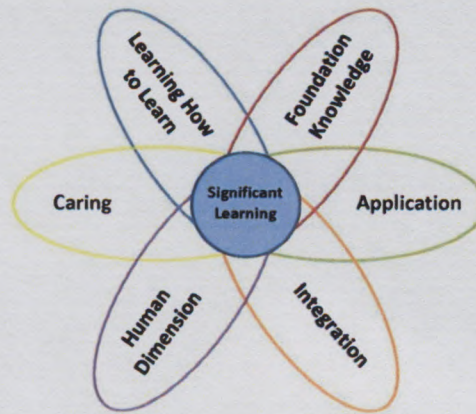


Figure 4- Fink's Taxonomy of Meaningful Learning. From Fink (2003).

2.3.2 Clinical Learning in Health Sciences Education

As already mentioned, the extent to which learning takes place should be maximised by effective teaching strategies. For this reason, educators within the healthcare domain are currently rethinking traditional teacher-centred teaching and learning and are developing new ideologies and philosophies to move towards a student-centred approach (Stanley & Dougherty, 2010). In the context of learning through experience, students are purposefully placed in clinical learning environments. These are: i) simulated environments during on-campus classroom practicals, ii) in-hospital rotations through different wards and areas and iii) pre-hospital clinical learning shifts on emergency vehicles.

There are numerous definitions for learning that occurs in authentic environments such as those listed above. Work Integrated Learning (WIL) is defined by the University of Johannesburg as the component of a learning programme that focuses on the meaningful application of

previously mastered theoretical understandings and practical competencies in an authentic learning work-based context under the supervision and/or mentorship of a person/s representing the workplace (University of Johannesburg, 2013). Clinical learning is listed as a specific activity that may take place during WIL. The definition that is offered by the CHE is shorter, serving more the function of an umbrella definition. The CHE regards WIL as a range of academic disciplines that integrate formal learning and workplace concerns. Clinical learning is regarded as a learning activity under the WIL umbrella term (CHE, 2011).

With the above taken into consideration, the author therefore sees clinical learning in this study as the learning that takes place when students are given an opportunity to interact with live patients in the authentic environment regardless of whether it is termed workplace-based learning, service learning, experiential learning and or internship.

2.3.3 Skills acquisition during clinical learning

Clinical skills acquisition occurs as students progress from one year to the next. A model of skills acquisition has been identified by Dreyfuss and Dreyfuss. The researcher identified the Dreyfuss model as a departure point to try and describe skills acquisition to ground later argument regarding competence in assessment. The reason for the use of this model as departure point is because other well-referenced authors have used the model as departure points for their models (Llyon, 2015; Benner, 94; Pena, 2010; Hall-Ellis & Grealy, 2013).

The Dreyfuss model was first published in 1980 by brothers Stuart E. and Herbert L. Dreyfuss, one an engineer and the other a psychologist (Dreyfuss and Dreyfuss, 1980). Within this model, the Dreyfuss brothers argue that humans undergo five stages when acquiring new skills. These

stages range from novice, the most basic stage of skill acquisition, to expert. Initially, the stages were i) novice, ii) competent, iii) proficient and iv) expert. Sometime later, the model was redefined and the following stages were identified: i) novice, ii) advanced beginner, iii) competent, iv) proficient and v) expert. The latter is the model that is currently used (Benson & Nixon, 2008; Pena, 2006).

2.4 ASSESSMENT IN HEALTH SCIENCES EDUCATION

Yorke (2012) argues that assessment is primarily done for three reasons. These are: i) the measurement and reporting of achievement, ii) promotion of learning and iii) as a method of quality assurance (Yorke;2012). As discussed earlier in this chapter, authors such as Kraftwohl, Bloom, Johnston and Cadorn, Bagnasco, Rocco & Sasso argue that appropriate assessment and feedback on assessment is crucial to meaningful learning. These arguments are supported by Keating, Dalton, & Davidson, who highlights that assessment not only reports on the level of performance associated with an assessment area, but also identifies areas for remediation and provides feedback to learners and module lecturers (Stein, 2010). Blom and Keating, Dalton, & Davidson also argue that, if the purpose of assessment is made clear, an accurate indication of required level of performance is given and opportunities for feedback is made known, preparation for assessments by students may drive or cause learning to occur (Blom, 2014; Keating, Dalton, & Davidson, 2009).

Earlier in this chapter, the researcher introduced to the reader the concept of 'meaningful learning'. The researcher showed that the authors that described meaningful learning made a case for accurate assessment thereof. Also shown was that one of the characteristics of meaningful learning is that it is developed through experience. The researcher then showed that in the context of this study, experiential learning, or the

learning that is developed through experience, occurs during clinical learning opportunities. If clinical learning is to be appropriately assessed, the researcher should introduce concepts, assumptions and principles that underpin assessment within the context of clinical learning, which is what the researcher does in this section. Afterwards, the researcher will elaborate on clinical assessment more, making a case for the importance of clinical performance assessment.

2.4.1 Assumptions and Principles that underpin assessment

As mentioned in Chapter one, assessment concepts include validity, reliability, credibility and feasibility. Bashook argues that high stakes assessments, such as those that the CPAI will be used for should include a well-engineered blueprint and evidence of validity and reliability that is credible to defend against challenges (Bashook, 2005). The South African Qualification Authority (SAQA) identified ten principles and assumptions of assessment (SAQA, 2015). The researcher grouped the assumptions and principles of assessment described by SAQA, Bashook, Prober, Yorke and the United Kingdom Postgraduate Medical Education Training Board Workplace based assessment Subcommittee into four main categories. These are validity, reliability, credibility and feasibility. The researcher will now present literature on these four concepts. The actual tests used to test these concepts in this research study and the defence of these tests are described in Chapter three.

Validity

Validity broadly refers to the ability of the assessment instrument to test what it was intended to test (SAQA, 2015; Salkind; 2004; Cook, Thomas & Beckman, 2006). Validity cannot be emphatically proven and neither is validation a once-off process. Also, no perfect assessment instrument

exists - therefore instruments should rather continuously be assessed and methods to improve validity should continuously be sought (Beckman & Cook, 2005). SAQA is explicit when stating that assessment methods need to be validated and refers to validation as activities undertaken to ensure validity (SAQA, 2015).

Although many different types of validity have been described, two validity themes have emerged when the researcher investigated assessment instruments in the health care education setting. These two forms of assessment were construct validity and content validity. SAQA also explicitly refers to these two forms of validity in the National Policy of Assessment (SAQA, 2015).

Construct validity is seen as the overarching theme of validity, and refers to the degree to which an assessment instrument is designed to measure specific outcomes (Rosenthal & Westen, 2003). Simply put, construct validity refers to whether the assessment instrument able to assess what it was designed to assess. The question that content validity is associated with is "can the instrument assess what it is supposed to assess?" Content validity refers to how well an assessment instrument is able to assess the assessment outcomes.

Reliability

Reliability refers to the ability of an instrument to yield the same results multiple times over. An instrument that yields valid results, but that is not capable to produce the same results when used under the same circumstances is of little value in the health care setting. There are many forms of reliability, and these forms may be assessed differently. The form of reliability that this study investigated was inter-rater reliability. Inter-rater reliability is also sometimes referred to as interobserver reliability or

agreement. Regardless of the term assigned to it, this form of reliability refers to the ability of the assessment instrument to yield similar results when used by more than one assessor to assess the same concept or student.

Inter-rater reliability refers to the extent to which two or more observers agree on a given outcome (Viera, Joanne and Garrett, 2005). For this study, Inter-rater reliability will refer to the extent to which the marks awarded by two observers regarding the clinical performance of a student agree. It could be argued that one should also assess test-retest reliability. The concern with test- retest reliability is that the test must be reproduced under the exact same circumstances over a period of time. Since the instrument will be tested in the real life setting, this would be impossible. One cannot reproduce, for example, a patient that has sustained a gunshot wound to the leg in the same setting, with the same respiratory and haemodynamic status and same ambulance crew in an uncontrolled environment.

Credibility

In essence, credibility refers to how well the assessment results are supported by users thereof (Bashook, 2005). The educators and students who will use the assessments need to be supportive of the assessment method and value its use. Prober argues that the most important issue of credibility is the issue of transparency. This argument is supported by SAQA (Prober, 2011; SAQA, 2015). In order for assessment to be transparent, students and educators need to know exactly what criteria students are being assessed on, and exactly how they will be awarded marks (Prober, 2011; SAQA, 2015). In addition, there must be trust that assessment will not be biased and that the criteria will be applied fairly, objectively and consistently (Prober, 2011; SAQA, 2015). The argument is

made that assessment needs to be based on direct observation, rather than indirect reporting (Prober, 2011).

Feasibility

Feasibility can be divided into the theoretical and practical problems of design, development, and production of an assessment method, as well as the administration, data analysis, reporting, and ongoing revisions and use of the method. In nearly all situations, feasibility becomes a question of available money, expertise, opportunity, resources, and time. The most efficient approach is to borrow a proven existing method, make minor changes to adapt it for use in the new setting, and hope that the method is as valid for the new setting and the different type of health provider as it was previously. This is the easiest approach, but doing borrowed assessments in a new setting or with different stimulus cases/items or different raters may not provide valid and reliable measures of the competencies the assessment was intended to measure (Joint Committee on Testing Practices, 1999; Linn, 1994). Creating custom forms should be sufficient to document a clinician's ratings of students and give feedback to students. That said, custom assessments may not be credible or defensible for important decisions relating to passing or failing a module without additional corroborative evidence of reliability and validity (Cadorin, Bagnasco, Rocco & Sasso, 2014). In addition to the question of validity and reliability, the researcher showed earlier in this chapter that there is a paucity of literature describing the use of clinical performance assessment instruments in South Africa. This further weakens the argument for borrowing a proven existing method of clinical performance assessment.

Practical concerns with using any assessment method, as noted above, are the time, expertise, and resources needed to use it properly and get

useful results. Most clinical settings lack one or more of these. Users, such as clinical assessors and students, reflect greater satisfaction with tests that are short, easy to administer and less costly (Cadorin, Bagnasco, Rocco & Sasso, 2014). The University is able to carry the cost of administration of the clinical performance assessment in the form of the CPAI. What needs to be established, though, is whether or not the time and effort it takes to complete the assessment is considered cumbersome. The investigation into the practical concerns with the feasibility of the CPAI will be dealt with in Chapters five (focus group discussion) and six (triangulation of data and discussion).

The researcher has presented literature that contextualises clinical assessment for this research study. The next section will be used to present literature on assessment methods to assess clinical performance in the international arena.

2.5 ASSESSMENT OF CLINICAL COMPETENCE

In this section, the researcher will present literature regarding competence and performance assessment, and make an argument for the need of a comprehensive clinical performance assessment strategy.

As mentioned in Chapter one, there is no opportunity for newly graduated emergency care practitioners to complete an internship programme upon graduation. This means that unlike medical school graduates, EMC students will not have the opportunity to “find their feet” during mentored post qualification internship. For this reason, universities must be responsible to ensure that the students they graduate from their EMC programmes are clinically competent and ready for independent practice upon graduation. The assessment of clinical competence is therefore regarded as an important and invaluable activity within the academic unit.

The conceptual reference that the researcher will use regarding the assessment of clinical competence is Millers Framework for assessment. Miller (1990) proposed a framework for assessment that is frequently used in health sciences education. Within his framework, Miller proposes four functional units, or levels, of assessment. These are arranged hierarchically beginning with assessment of simple content knowledge leading up to assessment of the meaningful application of learning to solve complex real world problems. Miller's four assessments of competence as applicable to the clinical learning and assessment domain are highlighted below:

Knows: This is considered the most basic assessment of knowledge and forms the base of the hierarchy. This type of assessment has reference to propositional knowledge, or "knows what" knowledge and is commonly assessed using simple written knowledge tasks.

Knows how: The ability to acquire information, to analyse and interpret information in order to formulate diagnostic or management plans based on the analysis and interpretation. The student's ability to think critically in a functional setting is assessed. This component may also be assessed by written knowledge tasks (tests and assignments) as well as viva voice assessments.

Shows how: Student behaviour is assessed at this, the third level. Evidence is gathered by observing student performance in standardised patient tasks. Standardised tasks is the key factor at this level, as the assessment is focused on the ability to perform tasks that is the same for all students and applications rather than pure cognitive function. Examples include patient simulations and observational structured clinical evaluations (OSCEs).

Does: The key phrase associated with the apex of the hierarchical framework is “action”. Assessment at this level involves supervisors or mentors observing performance in practice in a real world setting in lieu of a classroom setting. The CPAI will focus on the assessment of students in this area of Miller’s framework.

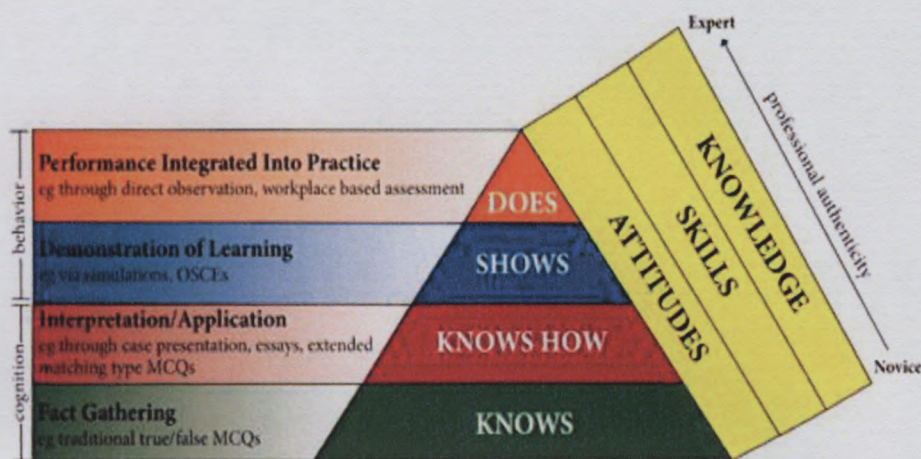


Figure 5- Miller’s Prism of Clinical Competence (Miller, 1990).

As can be seen from the above explanation, clinical competence is considered to be all of the activities related to the rendering of medical care to patients. Clinical competence is the culmination of classroom - based activities during theory lecture time, as well as practical assessments such as OSCEs, patient simulation assessments and clinical performance assessments. It includes all of the areas of Miller’s competence framework (explained below). Currently, ECP students at the UJ are already assessed using written knowledge tasks (tests, assignments, and case studies), Objective Structured Clinical Evaluations (OSCEs) and patient simulations.

Clinical performance is regarded as that which is measured or assessed during clinical learning. It refers to the “does” component of clinical

competence as proposed by Miller. Whilst the recognized theory exists, there remains a paucity of literature describing the use of performance assessment in the South African pre-hospital emergency care education setting.

Clinical performance assessments done in the authentic pre-hospital setting differ from classroom assessments because said performance assessments involve a move from what is known to what is unknown. Competence assessment is considered by some to be favourable as it is less resource intensive, easier to administer and more practical (Joint Committee on Testing Practices, 1999; Linn, 1994). Performance assessment is more resource intensive, it is more challenging to administer and potentially less practical because it is moved from a well-controlled classroom to an unknown, relatively uncontrolled real-world environment (Keating et al., 2009).

In the context of this study and within the UJ BEMS programme, clinical assessments are mostly linked to Clinical Practice Modules which are credit bearing and feature in all four years of the B EMC programme. To pass these module students have historically been required to see a set number of patients (caseload) and perform a minimum number of pre-determined clinical procedures. This caseload and clinical skill acquisition are requirements for promotion to the following year of study.

Historically however, these activities have not generated a mark. Rather, the mark for the clinical practice module was generated through the assessment of case studies, elective assignments and reflective journals. Although there is an area on the departmental patient report forms (which students must complete for each of the cases they have managed) that allows for the supervising clinician to provide feedback and comments on the clinical performance of the student, this is seldom utilised and what

little feedback is provided is often unstructured, brief and lacking in depth. Consequently this format has rightfully been criticized as it does not include an appropriate assessment of clinical performance in the authentic pre-hospital environment.

If this is to be remedied then educators would require a valid instrument for such an assessment. The investigation into instruments used to assess clinical performance in other health care domains informed the design and creation of the UJ CPAI.

Having mentioned the above, Keating et al. demonstrated that because the setting of the clinical performance assessment is so much different from the classroom settings, students do things differently in the real world setting compared to the classroom setting (Keating et al., 2009). For this reason there is general agreement that the need exists for a valid, reliable, feasible and credible assessment of what students actually do when they perform clinical learning shift rotations (Miller, 1990). This is the type of assessment that this study will focus on. In the next section, the researcher presents literature on methods employed to assess the clinical performance of students.

2.6 ASSESSMENT OF CLINICAL PERFORMANCE

As mentioned, clinical performance is best assessed in the real-world environment. A well- described method of clinical performance assessment is with the use of workplace-based assessment instruments (van der Veluten, 2011). These type instruments provide a snapshot (5-15minutes) of how students perform when they are performing duties in the working environment. These methods of assessment usually involve a number of competencies that are assessed and are therefore considered quite complex (SAQA, 2000; van der Veluten, 2011). SAQA, and van der Vleuten

regard workplace-based assessment instruments as unstandardised. A lot of literature is currently being published in the use of workplace-based assessment internationally. Van der Vleuten proposes that four lessons have been learnt through this work (van der Vleuten, 2011). These lessons are described below.

Workplace-based assessment may improve learning in the workplace. This is because adequately designed assessments allow assessors to provide good feedback to students. In-line with the learning philosophies that the researcher referred to earlier in this chapter, good feedback results in more meaningful learning (Dreyfuerst, 2012; PMETB workplace-based assessment subcommittee, 2005). Feedback does not have to be strictly quantitative – in fact, good feedback is considered to be quantitative combined with narrative (qualitative feedback) [van der Vleuten, 2011; PMETB workplace-based assessment subcommittee, 2005].

These types of assessments are not considered standardised (SAQA, 2000). Unlike standardised assessments, such as simulations, it is impossible to standardise the types of experiences, surroundings, other medical team members, available equipment, and assessing clinicians within the workplace-based environment. Although the CPAI will not change, the actual case / patient the student will be assessed on will differ. Because the assessments are not standardised, it may lead to subjectivity. This, however, may be overcome with sufficient sampling of clinical performance and assessors.

Training of teachers and staff is important (van der Vleuten, 2011; PMETB workplace-based assessment subcommittee, 2005). Assessor training will ensure that there is greater consistency of assessment, less bias, increased transparency (Bashook, 2005; PMETB Workplace based

assessment Subcommittee, 2005; Tavares et al., 2013). This will increase reliability as well as credibility of the assessment.

2.6.1 Workplace-based assessment instruments used to assess clinical performance

A desktop search revealed no published literature on the use of validated clinical assessment instruments in the South African pre-hospital context. Internationally, not many institutions include performance assessment as part of their assessment strategy. Some of the institutions that make use of performance assessment use pre-set instruments to guide the assessment of their students. These assessment instruments are often and referred to as Workplace-Based assessment instruments.

What the researcher found interesting, though, is that these workplace-based instruments were almost exclusively used for junior community service doctors as well as junior nurses and doctors soon after graduation. The researcher argues that this is perhaps in line with Dreyfuss and Dreyfuss when they argue that people are only really able to perform competently after they have gained some experience (Dreyfuss and Dreyfuss, 1980).

Although not a Workplace-Based assessment instrument, the researcher has identified the following competencies that are commonly mentioned in the issues of clinical performance. These are i) Patient care, ii) medical knowledge, iii) interpersonal and communication skills and iv) professionalism.

Tavares et al. sought to develop and validate a similar assessment instrument to the CPAI developed at UJ (Tavares et al., 2013a,b). This instrument assesses the clinical performance of students in seven

identified assessment areas: i) situational awareness, ii) history gathering, iii) patient assessment, iv) decision making, v) resource utilization, vi) communication and vii) procedural skills (Tavares et al., 2013b). This assessment instrument was the only instrument that the researcher could find that was developed specifically for use of the assessment of undergraduate students in the pre-hospital environment.

The Mini-Clinical Evaluation Exercise (mini-CEX) is a clinical instrument used by a number of institutions to assess medical doctors at the patient's bedside (Mitchell et al., 2011; Bouriscot et al. 2010; Moonen-van Loon et al. 2013). This short, 10-20 minute snapshot also consists of seven assessment areas. These are: medical interviewing skills, physical examination skills, humanistic qualities / professionalism, clinical judgement, counselling skills, organisation / efficiency and overall clinical competence (ABIM, 1995).

Mamelok (2009) investigated the use of workplace-based assessment portfolios for the licensing of medical practitioners in the United Kingdom. Twelve assessment areas were identified: i) communication and consultation skills, ii) practising holistically, iii) data gathering and interpretation, iv) making a diagnosis, v) clinical management, vi) managing medical complexity, vii) primary care admin and IMT, viii) working with colleagues and in teams, ix) community orientation, x) maintaining performance, learning and teaching, xi) maintaining an ethical approach and xii) fitness to practice. Mamelok reported that the instrument provided real opportunity for workplace-based assessment but there are weaknesses that require further development (Mamelok, 2009).

2.7 CONCLUSION

As mentioned in the beginning of this Chapter, this literature review was important to frame the study within the context of existing literature and knowledge. The researcher used this chapter to present literature on emergency care education in South Africa and the importance of clinical teaching, learning and assessment within the pre-hospital emergency care environment. The researcher also elaborated on the assessment strategies used to assess clinical competence and performance at the UJ Department EMC.

Prior to the development of the UJ CPAI, no formal way of assessing the performance students whilst they manage patients in the authentic pre-hospital setting existed. This was argued as problematic as both Rothans *et al.*, Miller and other authors highlight that decisions regarding clinical competence should not rely only on competence-based assessments in the classroom, but should include performance-based assessments in the real world setting. Their argument was that classroom-based assessments do not relate accurately to what occurs during clinical practice, and for this reason performance assessments are best done in the real world setting. In order to remedy the situation, the UJ CPAI was developed.

The researcher introduced literature on workplace-based assessment such as the CPAI, and showed how other institutions implemented workplace-based assessment instruments to assess clinical performance.

The researcher is in agreement with Mitchell *et al.* and argue that the implementation of the UJ CPAI as a new assessment strategy will only be successful if it is shown to be a valid, reliable, credible and feasible measure of clinical performance. For this reason, the central aim of this

study becomes the investigation into the validity, reliability and end-user support for the implementation of the CPAI.

The following Chapter will deal with the Research Methods employed in this study, as well as the Methodology used for the collection and interpretation of data and discussion thereof. Ethical considerations for this study will also be explained.

3.1 INTRODUCTION

In this chapter, the researcher reminds the reader as to the central aim, objectives and research paradigm applicable to this study which focused on exploring the validity and reliability of the UJ CPAI as an instrument for the assessment of clinical performance of emergency medical care students. This is followed by a detailed description and argument for the selected prospective, descriptive, sequential mixed method design. The chapter concludes with a discussion of ethical considerations applicable to the study and the methods followed to ensure validity, reliability and trustworthiness of the research.

In order to contextualise and defend the paradigm, choice of design and methods, the reader is now reminded of the central aims and objectives of this study.

3.2 AIM

As alluded to in Chapter the central aim of this study was to investigate and describe the validity and reliability of the UJ CPAI as an instrument for the assessment of clinical performance of emergency medical care students in the pre-hospital environment.

3.3 OBJECTIVES

In pursuit of the aim described above the following research objectives were determined:

- a) To investigate and describe (within the context of available data and existing literature) the structure and format of the UJ CPAI.
- b) To investigate and describe the scientific methods used to validate clinical performance assessment instruments and to select appropriate methods to validate the UJ CPAI in this study.
- c) To investigate and describe the construct and content validity of the UJ CPAI.
- d) To calculate and report on the Inter-rater reliability of assessment results using the UJ CPAI.
- e) Investigate and describe the views and opinions of educators regarding the UJ CPAI and its use for the the assessment of clinical performance of emergency medical care students in the pre-hospital environment and;
- f) To make recommendations regarding the structure, process, outcomes and future considerations for assessment of clinical practice within the UJ BHS EMC programme.

3.4 RESEARCH PARADIGM

The research paradigm used in this study is that of pragmatism. A research paradigm may be compared to the lens through which the researcher views reality. It provides an organised method of thinking, observing and interpreting findings within the setting of a specific discipline (Brink et al, 2012).

Scotland (2012) describes four components of a paradigm. These are ontology, epistemology, methodology and methods (Scotland, 2012). Within each paradigm, the researcher identifies a way of viewing reality (ontology) and knowledge (epistemology). The methodologies employed to investigate the research problem are dependent on the epistemological and ontological viewpoint of the researcher.

There are many different paradigms described in health care research. Followers of a positivist paradigm believe that reality is constant and that new objective knowledge may be produced if appropriate methodologies are followed. The methodology most commonly employed when following a positivist approach is quantitative in nature (Takavol & Zeinaloo, 2004)

Opposed to positivism is the paradigm of constructivism. Researchers following a constructivist paradigm see reality as subjective and meanings are constructed by people. Reality cannot be measured, as it is constructed by people based on experiences. The methodology followed is commonly qualitative.

There are limitations associated with pure quantitative and qualitative approaches. In an attempt to avoid these, a more recent paradigm, referred to as pragmatism, has been described in health care research. Here researchers have sought new methods to accurately investigate research questions and create new knowledge. Pragmatism draws on employing *what works best* when answering the research question. The pragmatic theory accepts that all human activity stems from the need to solve problems, and places emphasis in employing methods that are effective in order to best solve or answer the research problem (Andrew & Halcomb, 2009).

In this study the researcher adopted a pragmatic approach in order to answer the research questions. Through the lens of pragmatism, the researcher is able to value both objective as well as subjective knowledge (Morgan, 2007), which means that this paradigm is coherent with mixed method designs as described below (Morgan, 2007).

3.5 RESEARCH DESIGN

Research design is considered to be the logical action plan, or framework, which the researcher follows and uses to adequately connect gathered data and evidence in order to adequately address the research question (Yin, 1989; Flemke, 2004).

Quantitative designs are suitable for situations where the researcher has predetermined the options and more often a larger number of participants are involved (Anderson, 2006). The measurement in a quantitative design is objective and involves hard data and findings. Quantitative designs produce data that is statistically valid and generalizable to a population. The research is thought to generate reliable population-based findings which may be used to establish cause and effect relationships (Anderson, 2006).

In contrast to quantitative designs, qualitative research involves the collection, analysis and interpretation of non-quantitative data (Anderson, 2006). Qualitative research is often exploratory and open-ended using a variety of data collection tools to look for the subjective, opinions and feelings of participants. Qualitative research may employ a number of methods, such as analysis of literature and documents, interviews and focus group discussions. In qualitative research the emphasis is on rich, detailed and valid data in an exploration of that which is believed to be a dynamic reality. It is generally accepted that the knowledge produced by

qualitative research does not lend itself to be universal or replicable (Anderson; 2006).

A "newer" research design, mixed-methods research, has emerged where the researcher uses a combination of both quantitative and qualitative methodologies in a single study in order to answer the research question (Creswell & Plano Clark, 2011). Creswell and Plano Clark cite Johnson et al. (2004) as saying that a mixed method research design is utilized when the researcher wants depth and breadth of understanding and corroboration when conducting the research. Mixed methods design may be defined as the use of quantitative research to assess the magnitude and/ or frequency of constructs, and then employing qualitative research to explore the meaning and understanding of constructs using more than one method, and to then deliberately integrate these methods in order to draw on the strengths of each (Morgan, 2007; Andrew & Halcom, 2009).

Creswell argued that using a mixed methods design provided a better understanding of research problems than either a quantitative or qualitative approach on its own (Hesse-Biber, 2010; Creswell, 2009). It is also thought that by utilizing a mixed methods approach the researcher is able to obtain a more complete analysis of the problem (Creswell and Plano Clark, 2011). Creswell and Plano Clark felt that a mixture of the methodologies can assist in strengthening the reliability, validity and trustworthiness of the findings. These are discussed later in this chapter.

A prospective, descriptive, sequential mixed methods design was chosen for this study as this design allowed the researcher to value both qualitative as well as quantitative methods when attempting to answer the research question (Morgan, 2007; Hesse-Biber, 2010; Creswell, 2009). The study had to be prospective in nature as direct evidence did not exist in this area prior to this study. The sequential design was followed

because the researcher first needed to pilot the CPAI, gather and analyse the qualitative data and then using these outcomes to focus on the areas of qualitative data that were needed to fully answer the research questions.

3.6 RESEARCH METHODS

In order addresses the research objectives and achieve the central aim of the study the researcher needed to gather both quantitative as well as qualitative data in a logical, systematic fashion in the form of a prospective, descriptive, sequential mixed method. As the name of the design implies, the researcher therefore needed to make use of a number of research steps and methods. These are summarised in Figure 6 below.

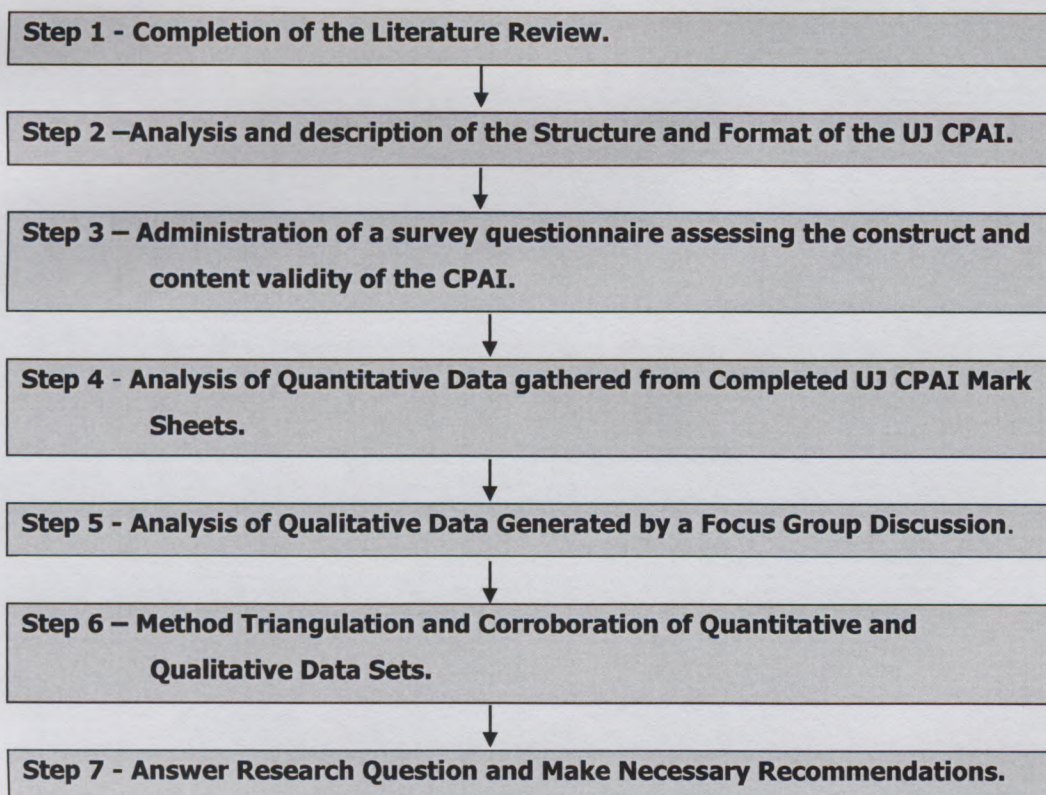


Figure 6- Research Methods used for this Study

Each of the seven research methods presented in Figure 6 above will now be described and their use defended in greater detail.

Step 1- Literature Review.

A literature review was done in order to contextualise the study. In the literature review (**Cf. Chapter 2**), the researcher investigated and described the purpose of the development of the CPAI. Theories that underpin learning and guide assessment were also presented and discussed in the literature review. The researcher also presented literature on the investigation and determination of validity and reliability.

Step 2- Analysis and description of the Design, Format and Structure of the UJ CPAI.

According to Beckman and Cook, no perfect assessment instrument exists. Instruments should therefore continuously be assessed and methods to improve validity should always be sought (Beckman & Cook, 2005).

The format and structure of the UJ CPAI was carefully assessed and described because the construct and content validity of assessment tools are directly affected by the structure and format thereof (Eden and Swoboda, 2013). As mentioned, the UJ CPAI is structured along nine assessment areas and the format of the UJ CPAI is a narrative informing a five-level rating scale per assessment area.

The structure and format of the UJ CPAI was described by seeking comparability of the assessment areas of the UJ CPAI and other internationally validated assessment instruments (**Cf. Chapter 4**). This was done by tabulating the different assessment areas of all of the

internationally validated instruments used in this study as well as the assessment areas of the CPAI. The researcher then determined whether the assessment areas that were used in the CPAI were similar to the internationally validated instruments. The researcher also manually counted how many assessment areas occurred in all of the instruments, and how many occurred in only one assessment instrument. The researcher then compared the scale used to award marks in the CPAI to the scales used to award marks used by the other assessment instruments. This was also done by tabulating the similarities and differences. The researcher also sought to find other published literature that validates the inclusion of the assessment areas in the UJ CPAI. The results are presented in Chapter four.

After having completed the abovementioned processes, the researcher was in a stronger position to carefully analyse and discuss later findings from both the quantitative and qualitative processes and to address issues such as construct validity, content validity and inter-rater reliability in Chapter 6 of this research report. Pertinent findings of the abovementioned investigation were also added to the discussion points for the focus group discussion.

Step 3 - Assessment of the validity of the CPAI in this research study.

Methods to assess validity

Construct validity may be demonstrated by comparing the ability of the assessment instrument to test what it was designed to assess against available literature or other validated tests (Salkind; 2004; Beckman & Cook; 2005). An example would be a new instrument designed to test for stroke. Typically, clinicians (or assessors) would be asked to assess a number of patients using the new instrument and diagnose the patients that have suffered a stroke. Then, the same clinicians would be asked to assess the patients again, this time using another validated instrument, such as the Los Angeles Prehospital Stroke Screen. The ability of the new assessment instrument to identify patients who suffered a stroke would then be compared to the ability of the validated instrument to do the same. Researchers may then make a finding of construct validity (Rosenthal & Westen, 2003). For this study such a direct comparison would not have been possible as no similar documented assessment instrument existed in South Africa at the time of this study. Also, the assessment areas from the other validated instruments differed from the UJ CPAI (**Cf. Chapter 4**) and therefore a formal comparison was not possible.

It is also described that construct validity may be qualitatively tested by arguing the relevance or necessity of the assessment areas the assessment instrument and against the construct being assessed. Another method of testing construct validity that is not as direct as the above methods, and is also not as widely described is a quantitative test that is used to determine consistency of ratings or marks awarded between different assessors or participants. Marks or ratings are either awarded by

answering questionnaires about the assessment instrument being assessed, or by using the assessment instrument for the purpose it was intended. Bashook as well as Rosenthal and Westen also argues that reliability is directly influenced by construct validity (Bashook, 2005; Rosenthal & Westen, 2003). Instruments that have poor construct validity is likely to show poor reliability.

A number of different methodologies to determine content validity have been described. A common method used is to make use of experts, who are required to rate the extent or depth to which an assessment is able to assess a given construct (Sireci & Faulkner-Bond 2014). A correlation coefficient is then calculated to determine to which extent the ratings awarded by each expert agree (Sireci & Faulkner-Bond, 2014).

Testing validity in this study

Construct validity was investigated both quantitatively as well as qualitatively. The construct validity was tested qualitatively by describing the structure and format of the UJ CPAI as described above.

Construct validity was also quantitatively assessed by having academic staff, clinical learning facilitators and external assessors complete a specifically constructed survey questionnaire (Annexure 6). This questionnaire is further discussed in the sections that follow. The researcher calculated the consistency of ratings awarded by the different participants that completed the survey by using a correlation coefficient. The correlation coefficient used for this study will be discussed later in this section. The results of this analysis are presented in chapter four.

Content validity was tested by calculating internal agreement of the responses provided in the questionnaire mentioned above, as well as by a

count of responses where ratings of four or more were awarded by the participants. Internal agreement was tested by calculating a correlation coefficient using one-way Analysis of Variance (ANOVA).

The Survey Questionnaire

The researcher invited nine prospective participants to complete a Self-designed, Non-validated, Evidence-informed Questionnaire assessing validity of the CPAI (Annexure 6). The questionnaire was based on published literature on validity evidence, where use was made of subject matter experts to rate the content validity of various objectives using a five-point Likert-type rating scale (Sireci & Faulkner-Bond, 2014). In addition, Copeland and Hewson made use of a questionnaire relating to the content validity of the Cleveland Clinic's Clinical Teaching Effectiveness Instrument (Copeland & Hewson, 2000). The questionnaire used for that study was similar to the questionnaire used in this study and asked participants to indicate the content validity of 15 items using a five-point Likert-type rating scale (Copeland & Hewson, 2000).

The questionnaire in this study required participants to rate the extent or depth to which they felt that the narrative of each of the nine assessment areas making up the UJ CPAI allowed them accurately score the student's performance. In other words, the participants were asked to provide their opinions about how well each assessment area of the CPAI is able to assess that which it is supposed to assess. Responses to the questions in the questionnaire were recorded using a Likert-type scale from one to five. A rating of one was considered to indicate poor content validity, and a rating of five was considered to indicate excellent content validity. For this study ratings of four or more were seen as indicative of good content validity. The descriptors for the Likert-type ratings are presented in Table 1. In order to allow the researcher to gather more valuable data, a free

response section was included in the questionnaire. This allowed for further, more in-depth data collection.

Rating	Description
1	The narrative does not enable the assessor to assess the area at all.
2	The narrative poorly enables the assessor to assess the area.
3	Neutral, or unsure.
4	The narrative enables the assessor to accurately assess the area.
5	The narrative enables the assessor to assess the area perfectly.

Table 1- Likert-type scale ratings and their descriptors

Participants who completed the survey questionnaire

The sampled population included all educators and clinicians involved in the supervision, teaching and assessment of students within the pre-hospital clinical domain. The sample for this study was purposefully chosen and included nine participants.

Inclusion criteria

Academic staff, clinical facilitators and external assessors that had used the CPAI to assess students on the UJ Primary Response Vehicle (PRV)¹ were all eligible to complete the questionnaire. The reason for this inclusion criterion was that these were the only individuals that would have used the CPAI in the setting it was intended to be used and

¹ A PRV is a small vehicle (such as a sedan) used to transport emergency care personnel and equipment to emergency incidents.

therefore had sufficient exposure, experience and insight into the instrument to make meaningful comment regarding its content validity.

Exclusion criteria

Academic staff, clinical facilitators and external assessors that had not used the CPAI to assess students on the UJ Primary Response Vehicle (PRV).

Description of the sample

The educators and clinicians that were identified for this study were four academic staff members from the UJ Department EMC, two clinical facilitators at the UJ department EMC and three external clinical assessors. External clinical assessors refer to clinicians who do not work for the UJ on a full-time basis, but who assists the Department EMC with clinical learning shift rotations on the Departmental PRV. External assessors were included as part of the sample because they were able to provide opinions from a different perspective to that of the academic staff and clinical assessors directly involved with the Department EMC. For this reason, using external assessors also improved the validity, reliability and trustworthiness of the study. The researcher and the supervisor were excluded from the study since they were too closely involved with the data. Of the nine experts invited to take part in the research, seven consented. These included four academic staff members at the UJ Department EMC, two clinical facilitators at the UJ department EMC and one external clinical assessor.

Analysis of responses to the questionnaire

The researcher analysed the data from the questionnaire in four ways as follows:

1. Testing consistency through calculation of a correlation coefficient using a two-way random model.
2. Testing for internal agreement by calculating Cronbach's Alpha using a one way ANOVA model.
3. Count of responses where ratings of four or more were awarded
4. Manual coding of free responses for those questions that allowed for such.

Each of the above methods is described in more detail below.

Calculating the Correlation coefficient to test consistency and internal agreement

The researcher tested construct validity and content validity by calculating a correlation coefficient. The researcher will now elaborate on correlation coefficients and then explain how consistency and internal agreement was calculated.

There are many different types of correlation coefficients such as Cronbach's Alpha, Kendall's Tau, Pearson's Correlation Coefficient (r) and Cohen's Cappa. The correlation coefficient measures the degree of agreement between variables, expressed as a numerical value (Salkind, 2004, Beckman & Cook, 2005). What this means is that the correlation coefficient provides a number that indicates to the researcher the degree or extent to which the participants awarded the same ratings when completing the questionnaire.

The correlation coefficient is expressed as a number that may range from -1 to 1. The closer the number is to zero, the weaker the agreement between the SMEs. A value close to 1 means that there was a strong agreement between the participants and a value of zero means that there is no relationship between the identified variables (Salkind, 2004, Beckman & Cook, 2005). Currently, with the advent of statistical software, the validity coefficient is calculated using software such as SPSS® (Salkind, 2004, Beckman & Cook, 2005). A correlation coefficient between 0.400 and 0.600 is considered to show moderate correlation, whereas 0.601 and higher shows high correlation between identified variables (Landis & Koch, 1977; Salkind, 2004).

For this study, Cronbach's Alpha was used. The reason was that, with this type of correlation coefficient, the researcher could easily instruct the software to calculate Cronbach's Alpha twice, each time doing so slightly differently (Norusis, 2003). More specifically, the researcher calculated Cronbach's Alpha using one way Analysis of Variance (ANOVA) to test internal agreement and as a two-way random model to test consistency (Norusis, 2003; Rosenthal & Westen, 2003; Landis & Koch, 1977; Salkind, 2004). The confidence interval for testing Cronbach's Alpha using a two-way random model was set at 95% and the cut-off value of 0.600 was identified to show consistency (Landis & Koch, 1977, Salkind, 2004).

Count of responses where ratings of four or more were awarded

The third test involved the researcher counting how many participants awarded rating of 4 or more for a given assessment area. The cut-off was identified by the researcher as 71%. This means that five of the seven respondents needed to award a rating of 4 or more to the assessment

area in order for the assessment area being deemed to have content validity. The results of this analysis are presented in the following chapter

Manual coding of free responses provided by the participants

The researcher collected, manually coded and interpreted the free responses provided by the participants during the completion of the questionnaire. Pertinent findings were added to the agenda of the focus group discussion for further, more in-depth discussion. The final comment regarding construct validity will be made in Chapter 6

Step 4 - Investigation of the Reliability of the CPAI by the Analysis of Results from Completed UJ CPAI Mark Sheets.

In order to determine the inter-rater reliability of the UJ CPAI, the instrument was introduced on the UJ Departmental Primary Response Vehicle (PRV) from February 2015 until November 2015. During this time, a research file was placed on the PRV. Inside this file was an information sheet on how to complete the UJ CPAI (Annexure 2), as well as a copy of the UJ CPAI narrative (Annexure 3) and a number of assessment forms (Annexure 4). Each of the students working a clinical learning shift on the UJ PRV were assessed independently by two clinical assessors. A total of 112 instrument mark sheets on 56 incidents were completed. The mark sheets for four incidents were spoilt because they were completed incorrectly and could not be used. The researcher therefore included 104 completed CPAI assessments from 52 incidents.

Students were assessed on the nine assessment outcomes, using a rating scale of one to five. The reader may refer to Annexure 3 for the assessment areas, as well as the narrative. The marks awarded for students by each assessor were captured per incident. The reader may

refer to Annexure 4 for the mark sheet. In other words, the two assessment mark sheets completed for a student on a specific call were captured together.

After that, the results were sent to a statistician in order to calculate a correlation coefficient for each assessment. A correlation coefficient simply expresses the degree of true agreement between a numbers of people (in this case, the two raters) as a numerical value. True agreement simply means agreement between people not due to chance. The correlation coefficient calculated for this study was Chronbach's Alpha, discussed above.

For this study, a coefficient of 0.60 or higher with a confidence interval of 95% was considered to show appropriate inter-rater reliability, and a value below that was considered to show sub-optimal inter-rater reliability (Norusis, 2003, Landis & Koch, 1977; Salkind, 2004). The researcher identified that 80% of the completed assessments needed to show appropriate inter-rater reliability in order for the researcher to make a recommendation regarding the inter-rater reliability of the instrument.

Step 5 - The focus group discussion.

A focus group interview was held to investigate and describe the views and opinions of educators regarding the CPAI and its use for the the assessment of clinical performance of emergency medical care students in the pre-hospital environment. Focus groups produce qualitative data and are considered valuable for researchers to assess opinions and ideas of participants (Skinner, 2007). Focus groups allow the researcher to explore certain concepts more in-depth, which is especially useful if there is not much literature on the topic. As mentioned in Chapter 2, literature on Clinical Practice assessment instruments in the international arena used

for the assessment of medical doctors and nurses were plentiful, but literature on validated pre-hospital clinical practice assessment instruments used in South Africa was lacking.

There was one focus group discussion. The researcher acknowledges that one focus group discussion does not lend itself to within method reliability; however, having a second focus group would not have been feasible. The reason for this is that the participants in the focus group had to be the same individuals that completed the questionnaire. Also, the discussion points were identified because of answers provided by the participants when completing the questionnaire. It would therefore not have contributed to within method reliability to have more than one focus group discussion.

The session was held in a booked boardroom on the premises of the University of Johannesburg Doornfontein Campus. In order to maintain data integrity and trustworthiness of the study, the interviews were recorded on audio file, and transcribed by the researcher, avoiding participant identifiers (Shenton, 2003). The structure and discussion points for the focus group discussion were carefully and purposefully designed according to issues raised in the literature search as well as the quantitative data findings.

Participants for the focus group discussion

Participants for the focus group discussion were purposefully selected and invited to take part (Annexure 7).

Inclusion criteria

The same academic staff members, clinical facilitators and external assessors at the UJ Department of Emergency Medical Care that completed the questionnaire were all eligible to take part in the focus group discussion. The rationale behind purposefully selecting these participants were that these were the individuals that used the CPAI to assess students and was therefore in a better position to provide thick, rich and meaningful data (Skinner, 2007; Shenton, 2004). Also, the structure for the focus group discussion was chosen because of answers provided by the participants when completing the questionnaire. The researcher would therefore be able to gain more in-depth understanding of the reasons for the ratings awarded in the questionnaire as well as more detailed answers on the other topics of the focus group discussion. Seven participants were in the focus group discussion.

Exclusion criteria

The academic staff members, clinical facilitators and external assessors that did not complete the questionnaire were not eligible to take part in the focus group discussion.

Discussion Points

As mentioned above, the discussion points were carefully and purposefully structured to investigate the views and opinions of educators regarding the CPAI and its use for the assessment of clinical performance of emergency medical care students in the pre-hospital environment. These points are attached as Annexure 10. The discussion points focused on of four focus areas. These areas were identified from the objectives of the study and analysis of the preceding processes. Under the heading of each

focus area, the researcher identified questions for discussion. These questions were identified from the pertinent findings of the discussion of the structure and format of the UJ CPAI, as well as the quantitative results stemming from the questionnaire as well as completed UJ CPAI mark sheets.

Analysis of the data

The focus group discussion was audio recorded. The audio recording was used to transcribe the focus group discussion verbatim. Two copies of the transcript were made. The one copy was then used to manually categorise the data using the 'cut and paste' technique (Krueger, 2015). Responses, themes and trends were then identified using thematic content analysis (Krueger, 2015).

Steps 6 and 7 - Triangulation of data, Answer Research Question and Make Necessary Recommendations.

Jick (1979) in Cassim (2014), defines triangulation as a combination of different methodologies in order to study the same phenomenon. Triangulation allows researchers to be more confident about their results and findings (Cassim, 2014). The type of triangulation used in this study is methods triangulation because the researcher searched for consistency of findings generated by different data collection methods (Patton, 1999).

As mentioned earlier in this Chapter, the main aim of this study was to investigate the validity and reliability of the newly developed CPAI. Each of the mentioned methods generated valuable data for this study. The final step in the research process was to combine the data gathered through the various methods. The pertinent findings were then integrated and

recommendations regarding the structure, process, outcomes and future use of the CPAI within the UJ BHS EMC Degree programme were made.

3.7 ETHICAL CONSIDERATIONS

Data for this study was gathered using feedback from staff members. Informed consent to participate in the study was obtained from all participants. The questionnaires and responses was treated as confidential. All data was stored in password-protected encrypted form on computer disk when not being used for analysis. The raw data was only available to the researcher and supervisors. During the write-up, it was not necessary to disclose the name of individual educators or students.

The researcher acknowledges the vulnerable nature of students being a part of a study. A number of steps have been taken in order to ensure that students participate purely out of individual choice. Final permission to invite students to participate in the study needed to be obtained from Professor Fourie from the Division for Institutional Planning, Evaluation and Monitoring (DIPEM) at the University of Johannesburg after the proposal had served at the Cape Peninsula University of Technology Faculty Research Committee, as well as the Ethics Committee. The ethics approval number was CPUT/HWS-REC2015/H06. Prior to the distribution of consent forms students received an invitation to participate in the study. It was emphasised that the decision to participate is voluntary, and that students may choose to withdraw from the study at any time without fearing any form of retribution. It was also made clear that participation in the study will not incur direct benefit to the student, academically (i.e. no extra marks was awarded), monetary or otherwise. Only students who consented for the study were allowed to participate. Decision not to participate did not result in retribution in any way.

The researcher also acknowledges the ethical concern of students being assessed whilst managing real patients. Whilst this could be considered undesirable, unfortunately no current feasible alternative exists for the assessment of pre-hospital competence in the workplace. As the research purpose was to enhance this assessment process, the concern abates.

3.8 VALIDITY, RELIABILITY AND TRUSTWORTHINESS OF THIS STUDY

The ultimate aim of research, regardless of the focus or methodology is to deliver findings that are valid, reproducible and trustworthy. As mentioned previously in this Chapter, validity refers to the extent to which a method measures what it is supposed to measure (Bryman, 2008). Reliability refers to the repeatability of a study i.e. where the same methodology applied to the same data would yield the same or similar findings. Trustworthiness, in addition to validity and reliability, refers to credibility and transferability (Shenton, 2004). The researcher will explain how validity, reliability and trustworthiness was ensured in this study.

The researcher appropriately used a prospective, descriptive, sequential mixed methods design for this study. Three different research methods were used that followed a logical sequence. These were a literature review, quantitative data collection and focus group discussion following the method mentioned above. That was followed by thorough method triangulation of findings. Trustworthiness of the study was ensured by careful identification of discussion points for the focus group discussion, audio-recording of the focus group discussion and transcribing the discussion verbatim. Also, the researcher carefully triangulated and discussed the data using a well described method called method triangulation.

The design of the questionnaire used was informed by evidence and literature on validity assessment as well as questionnaires used to validate other assessment instruments. The mark sheets of the CPAI were not edited or changed during the data collection process. The researcher was appropriately involved with the data collection process. The quantitative data was shared with a statistician from a service department of the University in order to calculate the various correlation coefficients mentioned earlier in this chapter. The quantitative data was therefore reviewed by an individual not involved with the researcher or the Department.

The points for discussion in the focus group were carefully identified from the pertinent findings of the quantitative data analysis. Audio-recording and verbatim transcripts of the focus group discussion were used to randomly spot check and confirm what was said by participants of the focus group discussion. Lastly, the quantitative and qualitative data sets were carefully triangulated by appropriate method triangulation and then discussed.

The transferability of this study is limited, because the researcher deliberately focused on the investigation of an instrument for the purposes of implementation at the Department of Emergency Medical Care at the University of Johannesburg. This study was not intended to be transferable or generalizable.

3.9 SUMMARY

This Chapter provided the background and context to this study which is to investigate the validity and reliability of a newly developed clinical performance assessment instrument (CPAI) at the University Of Johannesburg Department Of Emergency Medical Care. The research

problem, aim and objectives together with the mixed-methods sequential design and methodology were introduced.

The following Chapter focuses on presenting literature on assessment practices in the pre-hospital emergency care education setting, as well as the concept of determining validity and reliability.

CHAPTER 4- CONSTRUCT VALIDITY, CONTENT VALIDITY AND INTER-RATER RELIABILITY

INTRODUCTION

As mentioned in the preceding chapter the central aim of this study was to investigate the validity and reliability of the UJ CPAI as a tool for the assessment of clinical competence of Emergency Medical Care (EMC) students in the pre-hospital arena.

This Chapter will deal with the following:

- a) The analysis and description of the Structure and Format of the UJ CPAI (Research Step 2).
- b) The analysis of the quantitative data from the survey questionnaire which focused on the construct and content validity of the CPAI. (Research Step 3).
- c) Analysis of Quantitative Data gathered from the 104 Completed UJ CPAI Mark Sheets (Research step 4).

The Chapter is presented in four parts. The first part deals with a discussion on the format and structure of the CPAI. This is important because this explains the rationale for the inclusion of the different assessment areas, as well as how the rating categories work. Part one also allows for later interrogation, comment and argument around the construct and content validity of the CPAI as an assessment instrument in Chapter 6.

Parts two and three of this chapter deals with the construct and content validity of the CPAI respectively. This is done through presentation of the participant responses to the questions in the survey questionnaire.

The fourth and final part of Chapter four will focus on analysis of the quantitative data gathered from the completed UJ CPAI assessment sheets. This data was used to comment on the inter-rater reliability of the UJ CPAI.

4.1. PART ONE: FORMAT AND STRUCTURE OF THE UJ CPAI

As mentioned above, and in Chapter 3, the format and structure of any assessment instrument plays an important part of its validity. It is therefore important for this study to evaluate what the UJ CPAI assesses students on, as well as how students are rewarded.

This notion is supported by Edens and Swoboda, who in 2013 presented a 360-degree model for the development and implementation of workplace based assessment instruments (Edens and Swoboda, 2013). Their six stage approach is presented in Figure 7.

All six steps mentioned above have relevance in the development of the UJ CPAI. That said, the researcher argued that the steps that have the most relevance to the development of the UJ CPAI are *determining the areas that need to be assessed, determine specific behaviours to be evaluated for each area and determining an appropriate answer scale.*

Although the paper was presented after the development of the CPAI, the researcher saw the above four stages as a well-structured, logical method to investigate the structure and format of a tool such as the CPAI. This structure is explored below.

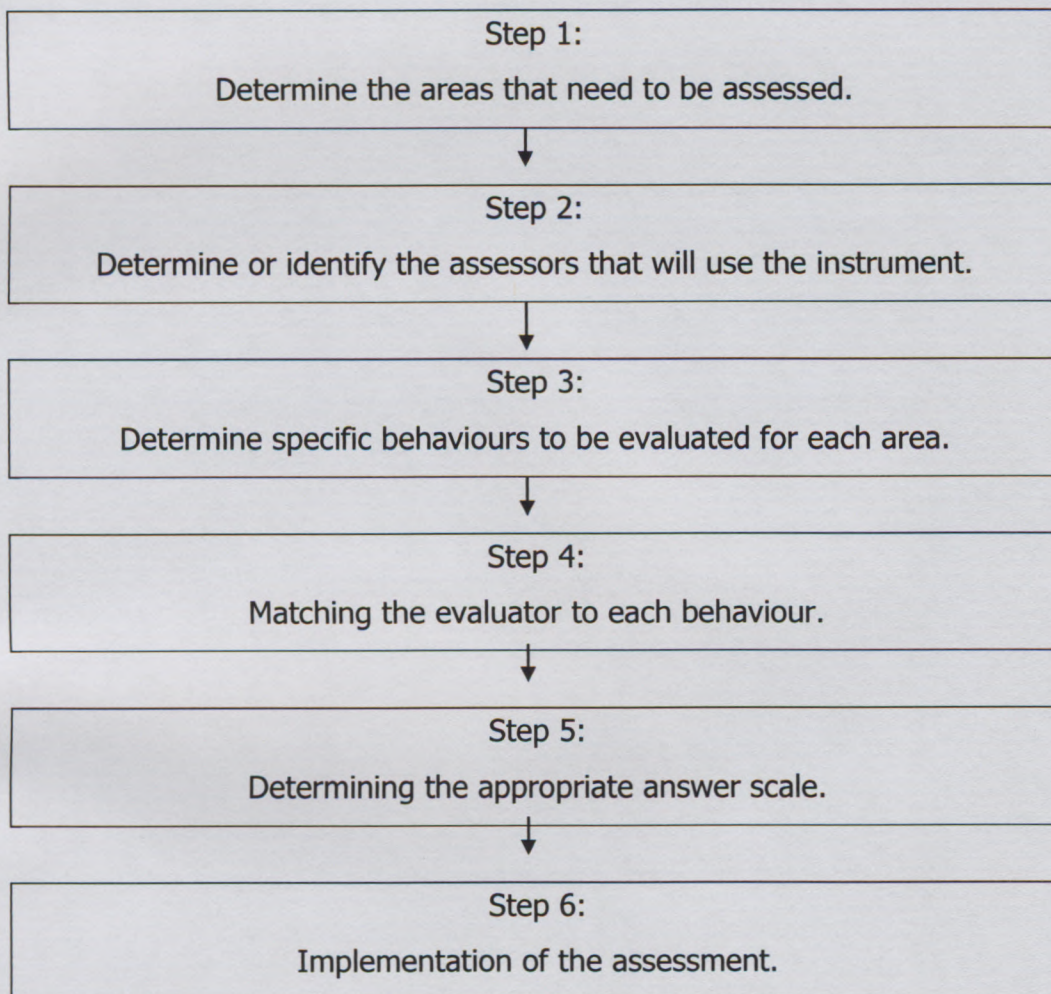


Figure 7- Eden and Swoboda's Six Step Approach to the Development of Workplace-Based Assessment Tools. (Adapted from Edens and Swoboda, 2013).

4.1 .1 Determine the areas that need to be assessed.

Identification of assessment areas is considered to be crucial in the development of any assessment instrument. For the CPAI, the identified assessment areas also provided the structure for the instrument. In the case of this study, the CPAI needs to assess clinical performance, i.e. what students do in the real-life setting.

The CPAI is structured along nine different assessment areas. Although one could argue that there are more than nine assessment areas that could be included, Yorke (2005) argues that the number of assessment areas influence the extent to which assessment tools areas may be accurately used (Yorke, 2005). Yorke shows that increasing the number of assessment areas increases the accuracy of an assessment instrument until a critical number is reached. After that, increasing the number of assessment areas in an assessment instrument decreases the accuracy thereof (Yorke, 2005). What that means is that one needs to include enough assessment areas in the CPAI in order to accurately report on the clinical performance of students. When too few, as well as too many assessment areas are included in a CPAI, the extent to which the instrument will accurately report on student clinical performance decreases. Yorke did however not indicate what the critical number was.

The reader is reminded that the nine identified assessment areas in the UJ CPAI are: i) Preparation, ii) Professional presentation, iii) History taking, communication and general patient interaction, iv) Clinical assessment and diagnostic skills, v) Decision making regarding clinical interventions, vi) Clinical procedures, skills and treatment interventions, vii) Clinical documentation and record keeping, viii) Handover and ix) Ability to reflect on own practice, self-critique and accept clinical direction. The rationale for including each of these assessment areas will be explained in more detail later.

4.1.1 .1 Comparability of the assessment areas of the CPAI and other assessment instruments.

The researcher compared the UJ CPAI to other validated CPAIs used internationally. Figure 8 represents this comparison.

When comparing the UJ CPAI to three other instruments (Fig 8), one can see that there are similarities in the areas being assessed. The assessment areas that were included in the majority of the assessment instruments reviewed were related to clinical assessment skills and diagnostic skills, procedural competence, clinical decision making and history taking and interviewing skills.

Interestingly, some areas that one would think would naturally form part of the assessment instruments, such as handover, were only included in the CPAI. The researcher argues that the reason for this is related to the setting in which the assessment takes place. Although all of the assessments are workplace based, practitioners and students will in some cases not be required to hand over their patient to other members of the health care team. An example of such an assessment is the Mini-Clinical Evaluation Exercise (mini-CEX), used in an in-hospital setting. This is important because the United Kingdom Postgraduate Medical and Training Board (PMETB) states that developers of workplace based assessment instruments must take the setting of assessment into account during the development of said tools. Assessment tools will only be valid if they are used within the setting they were designed for (PMETB, 2005:10-12).

Assessment area	UJ CPAI	Mini-CEX	UK WPBA	Tavares WBA
Ability to reflect on own practice, self-critique and accept clinical direction	✓			
Clinical Assessment and Diagnostic Skills	✓		✓	✓
Clinical Documentation and Record Keeping	✓		✓	
Clinical Procedures, Skills and Treatment interventions	✓		✓	✓
Community orientation			✓	
Counselling Skills		✓		
Decision making relating to clinical interventions	✓	✓	✓	✓
Efficiency		✓		
Fitness to practice			✓	
Handover	✓			
History taking, communication, general patient interaction	✓	✓	✓	✓
Humanistic Qualities / Professionalism		✓	✓	
Maintaining an ethical approach			✓	
Managing Medical Complexity			✓	
Overall Competence		✓		
Practicing Holistically				
Preparation	✓			
Professional Presentation	✓			
Resource Utilisation				✓
Situation awareness				✓
Working with colleagues and in teams			✓	

Figure 8- Comparability of the assessment areas included in the UJ CPAI, the Mini-CEX, UK WPBA and Tavares' WBA instrument.

Notwithstanding the difference in assessment areas in the different instruments, it is suggestive of a need for a purposefully designed CPAI with the South African undergraduate pre-hospital emergency care student in mind.

4.1.1 .2 **Comparability of the UJ CPAI assessment areas to the components of a pre-hospital incident.**

As mentioned above, the different CPAIs presented above focused on different assessment areas as those of the UJ CPAI. Since the areas assessed by the four instruments were different, the researcher argued that to comment on the structure of the UJ CPAI after having only done

the above comparison insufficient. For this reason, other, logical methods to investigate whether the nine assessment areas of the UJ CPAI were sufficient were sought. The researcher then tried to deconstruct ('break down') a routine pre-hospital incident (or call) into its different components and compared those to the assessment areas of the UJ CPAI.

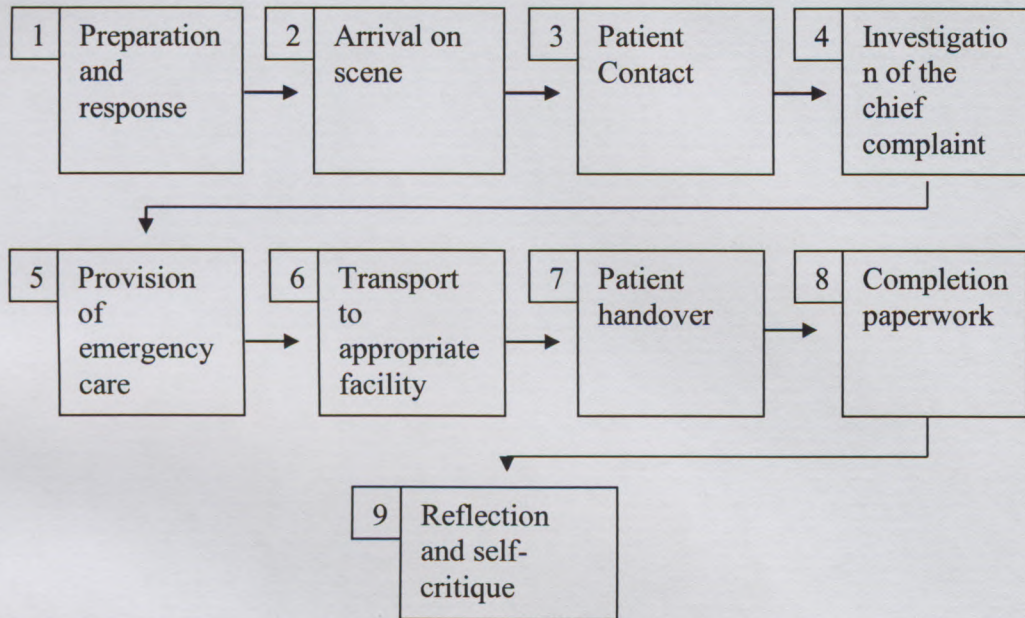


Figure 9- Deconstruction of a routine pre-hospital incident

As can be seen from the Figure, the nine assessment areas in the UJ CPAI are closely related to the different components of a pre-hospital incident. The researcher therefore argues that the nine different assessment areas included in the UJ CPAI are consistent with the components of a pre-hospital incident.

4.1.1 .3 Literature that validates the inclusion of the assessment areas in the CPAI

As mentioned earlier in this chapter, Edens and Swoboda argued that the assessment areas included in assessment instruments needed to be clearly

identified and carefully selected. Now that the nine assessment areas were identified, the researcher needed to investigate the rationale for selecting the assessment areas for inclusion into the UJ CPAI. The following section deals with the rationale behind selection of the different assessment areas and argues their value in this CPAI.

Assessment area one - Preparation

Preparation for the task at hand is arguably one of the most important duties the clinician may perform on a call-to-call basis (Sanders, 2001). One needs to prepare for the tasks at hand to ensure that procedures may be executed smoothly and that, when adverse event occur, equipment necessary to remedy the situation is ready and operational (Kovacs and Law, 2008).

Assessment area two - Professional presentation

It is important that health care workers present themselves in a professional matter. Professionalism is also an assessment area of the mini-CEX (ABIM, 1995). Although the Mini-CEX does not state that this refers to communication, respect and empathy amongst others, Faure (2014) cites the HPCSA when stating that professionalism not only refers to communication and behaviour, but also includes appearance (Faure, 2014). The Patient's Rights charter calls for the patient to be treated by a named health care practitioner (HPCSA; 2008). In addition to that, students need to dress with the appropriate personal protective attire as required by the policies of the Department EMC.

Assessment area three - History taking, communication and general patient interaction

Effective communication enables the researcher not only to obtain a comprehensive history, but it also builds rapport, establishes trust and creates a safe clinical environment (Bickley, 2003). The Batho Pele principles call for health care providers in South Africa to be courteous with customers of the health care system - i.e. patients (IPID, 2015). The patient's rights charter also declares that patients have the right to be treated by providers who display a positive disposition towards patients and that health care providers should demonstrate courtesy, human dignity, patience, empathy and tolerance (HPCSA, 2008).

The health history is considered one of three critical areas of clinical proficiency (Bickley, 2003). The health history, including the history of the presenting illness, forms the cornerstone of putting findings into context, making diagnoses and formulating management strategies (Bickley, 2003). In fact, the ability to obtain a comprehensive health history and communicate with patients effectively is one of only two assessment areas that were included in all four assessed workplace-based assessment tools (Tavares *et al.*, 2013b; ABIM, 1995; Mitchell *et al.*, 2011; Mamelok, 2009). Students need to repeatedly demonstrate the ability to communicate effectively with patients and obtain a thorough history prior to graduation.

Assessment area four - Clinical assessment and diagnostic skills

This is embedded in the UJ CPAI, UK WPBA and Tavares' WBA as necessary assessment areas (Tavares *et al.*, 2013b; Mamelok, 2009). Clinical assessment is considered the culmination of the health history (assessment area iii), the physical assessment (assessment area iv), and the written record (assessment area viii). The clinical assessment is

therefore aimed at not only assisting the clinician with diagnosis, but also to provide a record of what was done (Bickley, 2003).

Assessment area five - Decision making regarding clinical interventions

After the clinician has gathered enough information through the patient interview and physical assessment and developed a working diagnosis, the clinician needs to decide on applicable, valid, sustainable treatment plans (Bickley, 2003). The decision making process is a critical component to the overall management plan for the patient. Decision making ability is considered important enough to be the other of the two assessment areas included in all four assessment tools (Tavares et al., 2013b; ABIM, 1995; Mitchell *et al.*, 2011; Mamelok, 2009). The ability to make sound decisions based on the findings of the assessment process needs to be practiced. Clinical decision making is a process based on a number of steps that follow in a logical, sequential manner (Bickley, 2003).

Assessment area six - Clinical procedures, skills and treatment interventions

Part of the management plan for the ill or injured patient may include the performance of procedures or other interventions. Kovacs and Law argue that clinicians need to not only possess over cognitive and affective skills in competently managing ill and injured patients, but they need to have psychomotor ability as well (Kovacs and Law, 2008). Some of the indicated interventions may be performed by other members of the health care teams, but some complex procedures may only be performed by graduate degree practitioners (HPCSA, 2011). For this reason, it is important that the students need to demonstrate the procedural competence in a variety of procedures prior to graduation.

Although procedural competence may be demonstrated during OSCE practice in the classroom, assessors may find differently when procedures are performed in the real world environment (Keating et al., 2009; Bouriscot et al., 2010). Procedural competence has been identified as a necessary assessment area in the assessment tools developed by Tavares (2013) as well as the WPBA tool described by Mamelok (2009). The reason the Mini-CEX does not include procedural competency is because users of the Mini-CEX also use a DOPS (Direct Observation of Procedural Skills) assessment to assess the procedural competence of those students (Buriscot et al., 2010).

Assessment area seven - Clinical documentation and record keeping.

Stathers et al. argue that clinical documentation and record keeping is arguably the most important in the management of critically ill and injured patient (Stathers et al. 2011). Clinical documentation provides a record of the history, clinical findings of the physical assessment, vital signs, treatment plans, procedures executed and adverse events that occurred whilst the patient is in the care of the clinician and aids in continuity of care of the patient (Bickley, 2003).

In addition to the above, health care workers may need to rely on clinical documentation for legal purposes. For the above reasons, it is imperative to ensure that students and graduates are able to complete clinical documentation accurately and comprehensively prior to graduation.

Assessment area eight - Handover

As mentioned, once the patient arrives at an appropriate facility that is capable of providing more definitive care to the patient, or if the patient needs to be handed over to the care of another practitioner, a formal

handover procedure needs to take place. The purpose of the handover is to quickly relay important information that is necessary for the continuation of care to other practitioners in a well-structured, logical fashion. Handover is important for the continuity of care. Handover is only included in the CPAI as an assessment area.

Assessment area nine - Ability to reflect on own practice, self-critique and accept clinical direction.

As mentioned in Chapter 2, reflection and debrief by promoting conceptual understanding and supporting the development of clinical reasoning skills as a guide to professional practice (Dreyfuerst, 2012). Debriefing in Meaningful Learning (DML) has been well described. Dreyfuerst argues that debriefing during clinical learning allows students to reflect on their actions. This reflection is guided by teachers and mentors and encourages students to apply thought processes and knowledge to clinical experience. Debrief and reflection enhances their knowledge, decision-making skills and learning (Dreyfuerst, 2012).

Good reflection promotes coherence and cohesion between students' positive perception of the learning environment and practical proof of acquired learning (Dreifuerst 2012).

The researcher has identified that the structure of the UJ CPAI follows nine areas or assessment areas that are assessed in order to comment on the clinical performance of the students.

The format of the UJ CPAI is a narrative that informs a five-level rating scale. Assessors seek the narrative that best describes the performance demonstrated by the student for each assessment area. The rating scale attached to the narrative is then awarded to the student.

The next section will be dedicated to describing how the narrative written, and the five-point rating scale will be compared to the Dreyfuss model of skills acquisition. This will be done using steps three and five of the model by Eden and Swoboda presented earlier in this Chapter.

4.1 .2 Determine Specific "Behaviours" to be evaluated for Each Area.

In order to accurately assess students, care must be taken when behaviours are identified that need to be evaluated. In the case of the UJ CPAI, this meant that the narrative provided needs to explain clearly which behaviours, attitudes and competencies need to be demonstrated by students when performing clinical duties in the real world setting.

To this end, the researcher included the narrative of the UJ CPAI as Annexure 3 to this report. The reader is encouraged to read this narrative in order to see which behaviours, attitudes and competencies were included to assess each of the nine assessment areas.

4.1 .3 Determining the Appropriate Answer Scale.

Edens and Swoboda argue that the identification of an appropriate answer scale (i.e. *what* is regarded as competent, poor or exceptional performance) when engaged in the development of any WPBA instrument is an activity that should take significant effort (Edens and Swoboda, 2013).

A criticism by Stones (1994) regarding the definition of competence is that assessors using the instruments are not always sure exactly what 'competence' is. Yorke (2005) cites Efraut (2004) in saying that 'competence' means different things to different people. To some,

competence refers to social expectation. To others, it refers to personal attributes or qualities (Yorke, 2005). In order to define competence, Ashworth et al. suggest that it may be good practice to establish what practitioners actually do when engaged in clinical practice and then to attempt to describe those practices and attributes (Ashworth et al. 1994). Yorke indicates that another way of defining competence is to describe core competencies contained in the curriculum (Yorke, 2005). Even these approaches have been criticised by a number of authors, including Winter, Blake and Laing and Hussey and Smith (Yorke, 2005).

Because of the inconsistent definition of competence in the literature, each assessment area of the CPAI was supported by a narrative in the format of a five-point scale of assessment. When students are assessed, assessors identify the behaviours, attitudes and competencies demonstrated by the students. These are compared to the narrative or guide of assessment instruments in order to evaluate whether a *finding of competence* may be made. A finding of competence for this study means that the student has performed well enough to be awarded a mark of three on the marking scale for a specific assessment area.

The development of a narrative is supported by authors such as van der Vleuten (2011). The argument is that a rating scale with no or very little narrative does not provide adequate feedback, but may also be subject to bias and inconsistent awarding of marks (van der Vleuten, 2011).

The CPAI is the only instrument in this study that contained a narrative. The Mini-CEX and the workplace-based instrument designed by Tavares contained short definitions that supported each mark or term. Both the CPAI, as well as the workplace-based instrument designed by Tavares also included an area for qualitative feedback.

Comparability of marking scales

The marking scales of the different assessment instruments are presented in figures 10 – 13.

CPAI

Marking scale	1	2	3	4	5
Description	Does not comply	Needs improvement	Meets minimum standards	Exceeds Expectations	Far exceeds expectations

Figure 10- Marking scale of the UJ CPAI

Mini- CEX

Marking scale	1	2	3	4	5
Description	Insufficient evidence	Below expectations	Borderline	Meets expectations for completion	Above expectations

Figure 11-Marking scale of the Mini-CEX

UK WPBA

Marking scale	Not applicable	Not applicable	Not applicable
Description	Not met	Partially met	Fully met

Figure 12- Marking scale of the UK WPBA

Tavares' WPBA

Marking scale	1	2	3	4	5	6	7
Description	Unsafe	Unsatisfactory	Poor / Weak	Marginal	Competent	Highly competent	Exceptional

Figure 13- Marking scale of Tavares' WPBA

4.1 .4 Summary of Part One

During part one; the researcher described the format and structure of the UJ CPAI. The structure of the UJ CPAI follows nine assessment areas. The assessment areas of the UJ CPAI differ from other, similar assessment instruments used internationally. Despite this, all of the assessment areas follow a chronological structure similar to the components of a pre-hospital incident and are well described in the context of emergency medical care. It may therefore be argued that their inclusion to the UJ CPAI is logical.

The format of the UJ CPAI follows a narrative that informs a five-point rating scale for each assessment area. The narrative may be compared to the Dreyfuss' model of skills acquisition, as well as a logical description of what senior students at the Department EMC should be capable of when managing real-life patients. The five-point rating scale may also be compared to the Dreyfuss' model of skills acquisition, as this model also consists of five clearly defined rating categories.

The next part of this Chapter deals with the construct validity of the assessment instrument.

4.2. PART TWO: CONSTRUCT VALIDITY OF THE UJ CPAI

4.2.1 Overall ratings awarded by all the participants in the survey questionnaire.

As mentioned in Chapter 3, nine subject matter experts were invited to complete survey questionnaire about the content validity of the UJ CPAI.

Table 2 represents the ratings awarded by the seven participants across all nine assessment areas.

LICKERT SCALE RATINGS AWARDED TO THE ASSESMENT AREAS										
	1	2	3	4	5	6	7	8	9	Average all nine
Rater 1	3	4	4	2	4	3	4	4	4	3.6
Rater 2	4	4	4	4	4	4	4	4	4	4.0
Rater 3	5	4	5	4	5	4	4	4	4	4.3
Rater 4	3	3	4	4	4	4	2	4	2	3.3
Rater 5	5	5	3	4	4	5	5	5	4	4.4
Rater 6	4	3	2	4	4	4	4	4	4	3.7
Rater 7	4	4	3	2	2	2	4	4	4	3.2
Average all raters	4.00	3.86	3.57	3.43	3.86	3.71	3.86	4.14	3.71	

Table 2- Ratings awarded by all seven participants for all nine assessment areas

Out of a possible 63 ratings that could be awarded by the participants, a rating of four or five was awarded 78% of the time (n=49). Of the 49 times, a score of 4 was awarded 41 times and a score of 5 was awarded 8 times. The average scores awarded per assessment area ranged between 3.43 and 4.14 (mean 3.79, median 3.86). The average ratings awarded across the nine assessment areas per participant ranged between 3.2 and 4.4 (mean 3.8, median 3.7). These results will be discussed in more depth in part three.

4.2.2 An Analysis of the Consistency of the Results.

The two-way random Cronbach's Alpha for this questionnaire was -0.788 (95% CI, upper bound -0.901, lower bound -0.487). What this shows is that the ratings awarded by the participants were inconsistent. There may be a number of reasons for inconsistent ratings. The most likely reason is poor (or at least sub-optimal) construct validity. The researcher will discuss both of the above reasons in Chapter six. Once the consistency was calculated, the researcher carefully looked at the ratings awarded by the participants to see if reasons or examples for the low Cronbach's Alpha could be found. The prevalent findings will be reported on now.

The only assessment area with consistent ratings was assessment area 8. All of the seven participants awarded high ratings for the assessment area.

	P 1	P 2	P 3	P 4	P 5	P 6	P 7
Area 8	4	4	4	4	5	4	4

Table 3 - Ratings awarded to Assessment area 8 by participants

All of the other assessment areas did not show consistent ratings. The areas with the most inconsistent ratings awarded were areas two, three, four and six. When referring to table 3, the reader will notice that in every one of these areas, participants awarded a number of high ratings, as well as low ratings. The researcher uses area three as an example: Area three was awarded high ratings four times, and low ratings three times. What the researcher found inconsistent is that one participant (P 3) awarded the highest rating possible, yet another participant (P 6) awarded a rating on the complete opposite side of the rating scale. The same is applicable to Area six (participants 5 and 7).

	P 1	P 2	P 3	P 4	P 5	P 6	P 7
Area 2	4	4	4	3	5	3	4
Area 3	4	4	5	4	3	2	3
Area 4	2	4	4	4	4	4	2
Area 6	3	4	4	4	5	4	2

Table 4- Ratings awarded by the participants to assessment areas two, three, four and six.

When the researcher compared the participants to each other, it became evident that, in some cases, participants awarded almost opposite ratings for each other. Table 5 represents participants four and seven. Both awarded similar average scores (3.3 and 3.2 respectively). This means they have awarded similar numbers of high as well as low scores. The interesting finding was where they awarded the scores. In five of the nine assessment areas, the marks awarded by these participants were exactly the opposite. Where one participant awarded a high rating, the other awarded a low rating and vice versa.

	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8	A 9	Average
Participant 4	3	3	4	4	4	4	2	4	2	3.3
Participant 7	4	4	3	2	2	2	4	4	4	3.2

Table 5- A Comparison of Ratings Awarded by participants four and seven. A = Assessment Area

The findings of the quantitative data described above will be discussed in Chapter 6 when the data sets are corroborated.

4.2.3 General Comments on the Free Writing Component of the Questionnaire

After the section that allowed participants to award the Likert-type ratings to the content validity of the UJ CPAI, there was a section where participants were allowed to provide free responses should they so wish. A space was provided for the participant to write down the assessment area they wish to comment about and the rating awarded for the assessment area. The space to provide the free comment followed below that. These free responses were captured by the researcher. The responses were then grouped per assessment area and manually coded for trends. The trends identified by the mentioned process are presented in next.

Assessment area 1 – Preparation

A theme emerging from this section was that it was difficult to assess students on subsequent incidents after being marked the first time.

"The narrative needs to be expanded to accommodate more than on call/subsequent calls from the start of shift. The narrative only mentions start of shift preparation and does not include during shift and after call preparation, i.e. cleaning and restocking. Difficult to mark student when they are working as partners." - Clinical Assessor

"This can / should only be assessed at the start of the shift. Some criterion (sic) are not applicable to every call." – Lecturer

Assessment Area 3 – History taking, communication and general patient interaction

There was a general feeling that the narrative for this section was too vague.

"The biggest problem here is putting history taking together with the others. It is possible to *communicate* and *interact* well, but take a poor history and vice versa. This then pools too many unrelated things and makes it difficult to differentiate." - Lecturer

Assessment Area 6 – Clinical Procedures, skills and treatment interventions

It was noted that assessors found it difficult to award marks for this section if the student worked together with a partner.

"Students work with partners and need to rotate per call, but also when working on a patient, multiple procedures are performed by many people. Maybe criteria / narrative must change to describe if student correctly oversees the procedures and does effectively plan ahead, and effectively corrects partner / ems when something is done incorrectly." - Clinical Facilitator

"Although the criteria speak to the skills the student needs to possess, it is difficult to assess when students assist each other on a call or assessment of a patient. Ideally one student should perform all skills, preparation decisions etc." – Clinical Facilitator

General Comments

Some of the SMEs provided general comment that was not aimed at addressing a specific assessment area. These were also grouped and coded. There was a strong feeling amongst the participants that a five-point rating scale did not seem to work in the UJ CPAI.

"Item 5 seems detached from all of the criteria because it is so completely worded. There is no progression from 1-4-5 in any of the criteria." – Lecturer

"The narrative for a score of 4 already denotes perfect practice thus making the score of 5 null and void." - Clinical facilitator

"There is very little congruence between 1 and 5. I would assume that 1 would be the opposite of 5, or the negative, however, I struggled to find the links between the best and the worst. It would be better if 1 was a poorly performed 5, currently the two are too dissimilar." – Clinical Facilitator

4.2.4 Summary of Part Two

The researcher presented the results of the data analysis to test construct validity above. The results showed that the two-way random Cronbach's Alpha was low. This means that the ratings provided by the Subject Matter Experts were inconsistent. Also, in line with the poor consistency, it was found that, in some cases, participants awarded the exact opposite marks for the same assessment areas. This shows a high likelihood of sub-optimal construct validity. What this means is that it is likely that other tests, especially inter-rater reliability will also be sub-optimal. The researcher carefully looked at these findings and will discuss them when all of the data sets are corroborated.

The free responses that were generated from the open section of the questionnaire lead to the identification of a number of themes that were included in the agenda for the focus group discussion for further investigation.

As mentioned earlier in this Chapter, as well as Chapter three, this study did not only investigate the construct validity of the UJ CPAI, but also the content validity. In part three, the quantitative results stemming from tests performed on the content validity will be presented.

4.3 PART THREE: CONTENT VALIDITY OF THE UJ CPAI

4.3.1. Summary of the Ratings Awarded by the participants

The researcher presents table 6. Out of a possible 63 ratings that could be awarded by the participants, a rating of four or five was awarded 78% of the time (n=49). Of the 49 times, a score of 4 was awarded 41 times and a score of 5 was awarded 8 times. The average scores awarded per assessment area ranged between 3.43 and 4.14 (mean 3.79, median 3.86). The average ratings awarded across the nine assessment areas per participant ranged between 3.2 and 4.4 (mean 3.8, median 3.7). The one-way ANOVA Cronbach's Alpha was 0.752 (CI 95%, upper bound 0.911, lower bound 0.383). This means that, overall, the participants agreed that the UJ CPAI had good content validity.

LICKERT SCALE RATINGS AWARDED TO THE ASSESMENT AREAS										
	1	2	3	4	5	6	7	8	9	Average all nine
Rater 1	3	4	4	2	4	3	4	4	4	3.6
Rater 2	4	4	4	4	4	4	4	4	4	4.0
Rater 3	5	4	5	4	5	4	4	4	4	4.3
Rater 4	3	3	4	4	4	4	2	4	2	3.3
Rater 5	5	5	3	4	4	5	5	5	4	4.4
Rater 6	4	3	2	4	4	4	4	4	4	3.7
Rater 7	4	4	3	2	2	2	4	4	4	3.2
Average all raters	4.00	3.86	3.57	3.43	3.86	3.71	3.86	4.14	3.71	

Table 6- Ratings awarded by all seven assessors for all nine assessment areas.

4.3.2. A Comparison of the Ratings Awarded for Each Assessment Area

When one looks at tables 7 and 8, one can see that assessment area 8 was awarded the best ratings. All of the seven participants agreed that assessment area eight had good content validity. 86% (n=6) of the participants agreed that assessment areas seven and nine had good content validity, followed by assessment areas one, two, four and six with 71% (n=5) of participants awarding a good rating. Assessment area 3 was awarded the least number of good ratings. 43% (n=3) agreed that this assessment area had good content validity.

Number of good scores awarded per assessment area versus number of poor scores									
Assessment area	1	2	3	4	5	6	7	8	9
Score of 4 or 5	5	5	3	5	6	5	6	7	6
% Scored 4 or 5	71	71	43	71	86	71	86	100	86
% scored below 4	29	29	57	29	14	29	14	0	14
Total	100	100	100	100	100	100	100	100	100

Table 7- Number of Good Ratings Awarded per Assessment Area

Table 7 represents a comparison between the numbers of good ratings awarded per assessment area. Again, one can see that assessment area 8 received the highest number of good ratings. This area was the only to be awarded a good score by every participant. The assessment area with the least number of good ratings was assessment area three, with three good ratings awarded.

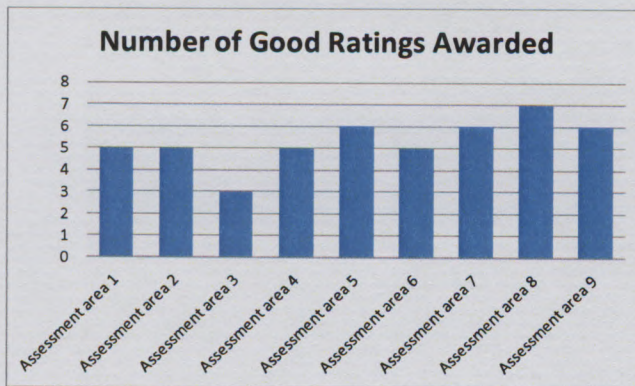


Figure 14- Comparison of Assessment Areas According to the Number of Good Ratings

4.3.3. Summary of Part Three

In part 4.3, the researcher presented another set of quantitative data. This data was generated by the completion of a survey questionnaire. Seven participants agreed to complete the questionnaire. Eight of the nine assessment areas were awarded good ratings by at least five participants. Assessment area three was awarded a good score by only participants. Assessment area eight was awarded the best ratings, and assessment area three was awarded the worst ratings.

The next section of this Chapter will present the second set of quantitative data that investigated the inter-rater reliability of the UJ CPAI.

4.4 PART FOUR: INTER- RATER RELIABILITY OF THE UJ CPAI

4.4.1 Results from the UJ CPAI mark sheets

As mentioned in Chapter 3, the CPAI was used on the UJ Departmental PRV from February 2015 until November 2015. During this time, a research file was placed on the PRV. Inside this file was a copy of the UJ

CPAI narrative (Annexure 3) as well as a number of assessment forms (Annexure 4). Each of the students working a clinical learning shift on the UJ PRV were assessed independently by two clinical assessors. This means that *one student* was observed and assessed by *two assessors* while managing *a specific patient* or case.

For each case, the student was assessed on nine assessment areas, on a scale of one to five. The reader is encouraged to see Annexure four for the UJ CPAI mark sheet. The student was assessed by the two assessors independently. The marks sheets were then placed back in the allocated compartment in the research file. The researcher collected all of the mark sheets after the completion of each shift and captured the data in a Microsoft Excel® spreadsheet, as seen below in table 4.4.1. This spreadsheet was shared with a statistician in order to calculate Cronbach's Alpha (α), using one way ANOVA. This allowed the researcher to comment on agreement (one way ANOVA). A 95% confidence interval was used for all 52 cases. In order to not burden the reader with too many pages of tables in this Chapter, the results of the statistical analyses is presented in Annexure 11.

The researcher was not able to calculate the inter-rater reliability across all nine assessment areas for all 52 cases, as the students, assessors and scenarios were not all the same for each of the 52 cases.

4.4.2 The Frequency of Cronbach's Alpha of More than 0.6

After calculating α for all 52 cases, the researcher then calculated the number of times α of 0.60 or higher was found. This data will represent the results of the one-way ANOVA only, as this is directly related to the assessment of inter-rater reliability. Out of the 52 cases completed, α of 0.60 or higher was found 35 times. Figure 15 shows the number of times

α of 0.60 or more was found, compared to the number of times a smaller number was found.

As can be seen from the data presented in Figure 15, below, a good α (≥ 0.60) was found in 67% of the cases ($n=35$), compared to 33% of the cases where a "poor" correlation coefficient (< 0.60) was found ($n=17$).

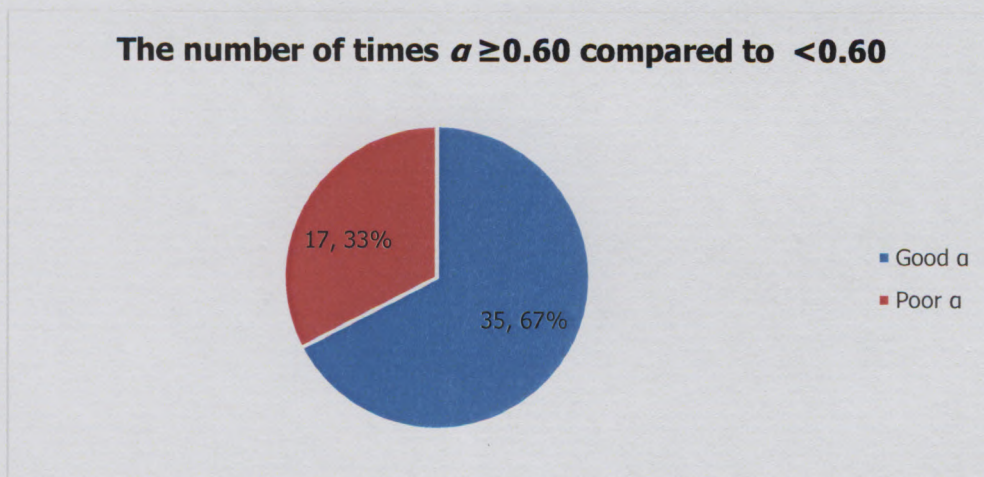


Figure 15- Number of times a Cronbach's Alpha of 0.60 or higher was found, compared to the number of times a smaller coefficient was found.

The researcher then grouped α according to intervals, as can be seen in Figure 16. The first interval was less than 0, after which the intervals were in increments of 0.1. The number of times α was found for each interval was then calculated.

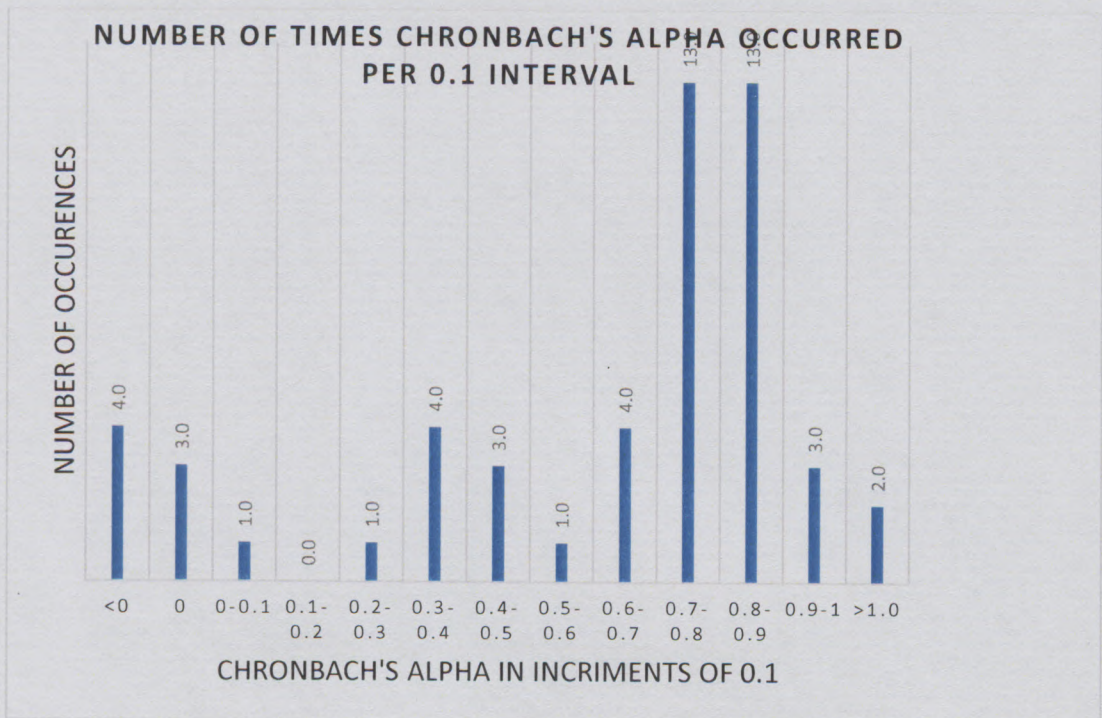


Figure 16- Graphical representation of each Correlation Coefficient group

An α of between 0.60 and 0.69 occurred in 7% of the cases ($n=4$). In the 0.7-0.79 as well as the 0.8-0.89 brackets, *the number of occurrences is 25% each ($n=13$)*. This means that 50% of the cases ($n=30$) fell within two intervals (0.70-0.79 and 0.80-0.89 respectively). This, however, was offset by a very low or inverse inverse α (less than 0) occurring in 15% of the cases ($n=8$). The average α across all 52 cases was 0.668 (mean – 0.668, median – 0.750, mode - 0.000). It happened twice that an α of more than 1.0 was found. *This is an anomaly. The researcher disregarded these two assessments.*

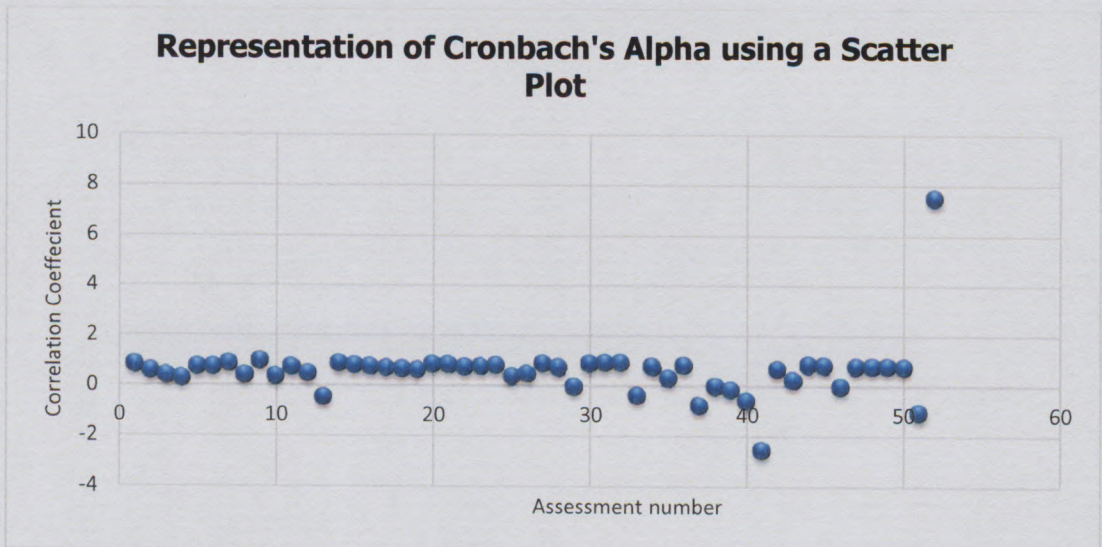


Figure 17- A representation of the Cronbach’s Alpha for each Case Completed

The above scatter plot shows that the majority of the coefficients were within the same range. This means that there is some data to support that the UJ CPAI is already able to produce reliable results. This made for an excellent question to add to the focus group discussion.

4.4.3 Inter-rater Reliability per Assessment Area

After the researcher analysed the data for a case-by-case basis, the researcher then looked at the inter-rater agreement within each assessment area. The researcher first calculated the correlation coefficient for each assessment area across all 52 cases, and thereafter calculated the number of times the two assessors awarded the same mark versus the number of times they did not award the same mark. The results are presented in table 8.

	Area1	Area2	Area3	Area4	Area5	Area6	Area7	Area8	Area9
Agree	26	32	24	26	25	22	30	29	25
Disagree	21	19	25	25	22	24	22	23	24
Incomplete	5	1	3	1	5	6	0	0	3
Percentage	55	63	49	51	53	48	58	56	51
Total	52	52	52	52	52	52	52	52	52

Table 8- The Number of Cases Where Assessors Agreed

As one can see from the table, assessors awarded the same marks 32 out of 51 times (63%) in Assessment Area 2. The one case where either one or both assessors did not complete the mark sheet was excluded. This was the most number of times the assessors awarded the same marks. The assessment area where the least number of instances were where assessors agreed was assessment area six. The percentages where the assessors agreed ranged from 48% (n=22/46) to 63% (n=32/51), mean 53.72%, median 53.19%. In each case, the researcher compensated for the number of times either one or both assessors did not complete the section.

The researcher then calculated a simple correlation coefficient (r) for each assessment area across all 52 cases. This was done using Microsoft Excel®. One should be careful to *directly* link the inter-rater reliability for each assessment area based on the correlation coefficient for 52 cases. The reason for this is because it was not the same assessors that used the CPAI 52 times. The researcher, did, however calculate the correlation coefficient for each assessment area in order to ascertain whether there was general agreement on the assessment area. The argument for this is that, if it happened often that the assessors awarded the same mark to a student for the specific assessment area, r would be high, and vice versa. Calculating r also allowed the researcher to comment on the degree or level of agreement. This means that, in addition to simply reporting on

how many times the same marks were awarded, the researcher could indicate how many times similar marks were awarded versus very dissimilar marks. A hypothetical situation is presented in table 9. This table represents two assessment areas. With both assessment areas, raters were asked to assess students 10 times. The number of times the same mark was awarded ($n=6$) is the same for both assessment areas. However, in the four cases where the assessors awarded different marks, the marks were similar in assessment area 1 and dissimilar in assessment area 2. Therefore, r for assessment area 1 was higher than for assessment area 2 (0.88 vs. 0.48). One can therefore argue that assessment area 2 warrant greater investigation and change.

Assessment Area 1				Assessment Area 2			
	Assessor 1	Assessor 2	r		Assessor 1	Assessor 2	r
Case 1	3	3		Case 1	3	3	
Case 2	2	2		Case 2	2	2	
Case 3	4	4		Case 3	4	4	
Case 4	5	5		Case 4	5	5	
Case 5	1	1		Case 5	1	1	
Case 6	2	2		Case 6	2	2	
Case 7	2	3		Case 7	2	4	
Case 8	3	4		Case 8	3	5	
Case 9	3	4		Case 9	1	4	
Case 10	5	4		Case 10	3	1	
r			0.88	r			0.48

Table 9- A Hypothetical Scenario Demonstrating the Value of Using r in Calculating Inter-Rater Reliability per Assessment Area.

The majority (8 out of the nine) of the assessment areas showed similar r to percentage agreement. The only assessment area where there was a significant discrepancy between r and percentage times the same mark was awarded was assessment area 2. Assessment area 2 showed the highest percentage times where the same marks were awarded (63%, $n=32/51$), yet showed the lowest r (-0.37). Figure 18 represents a

comparison of the percentage times the same mark was awarded versus the correlation coefficient. Series one represents r and series two represents the percentage of times the same mark was awarded.

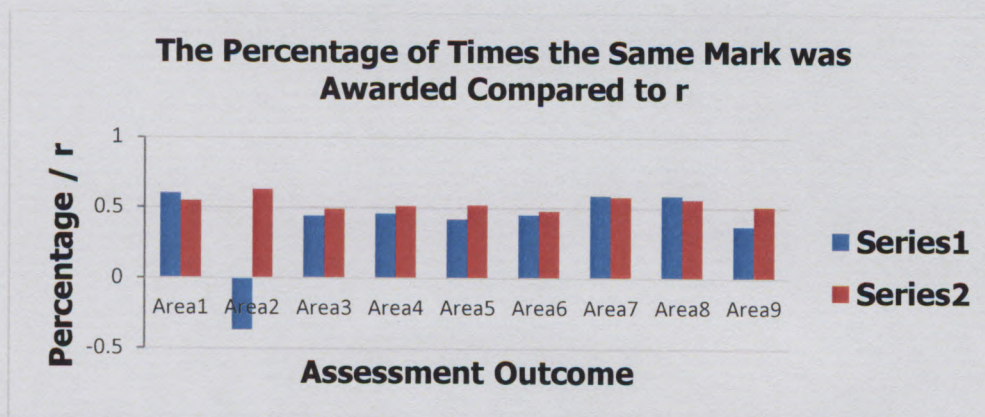


Figure 18- Correlation Coefficient for each Assessment Area compared to the percentage of times the same mark was awarded.

4.4.4 Comparison of Cronbach's Alpha using one Way ANOVA and two way Random Models

Finally, the researcher compared the one-way ANOVA σ with a two-way random α for the 52 cases. This was done to compare the absolute agreement calculated by one way ANOVA to consistency.

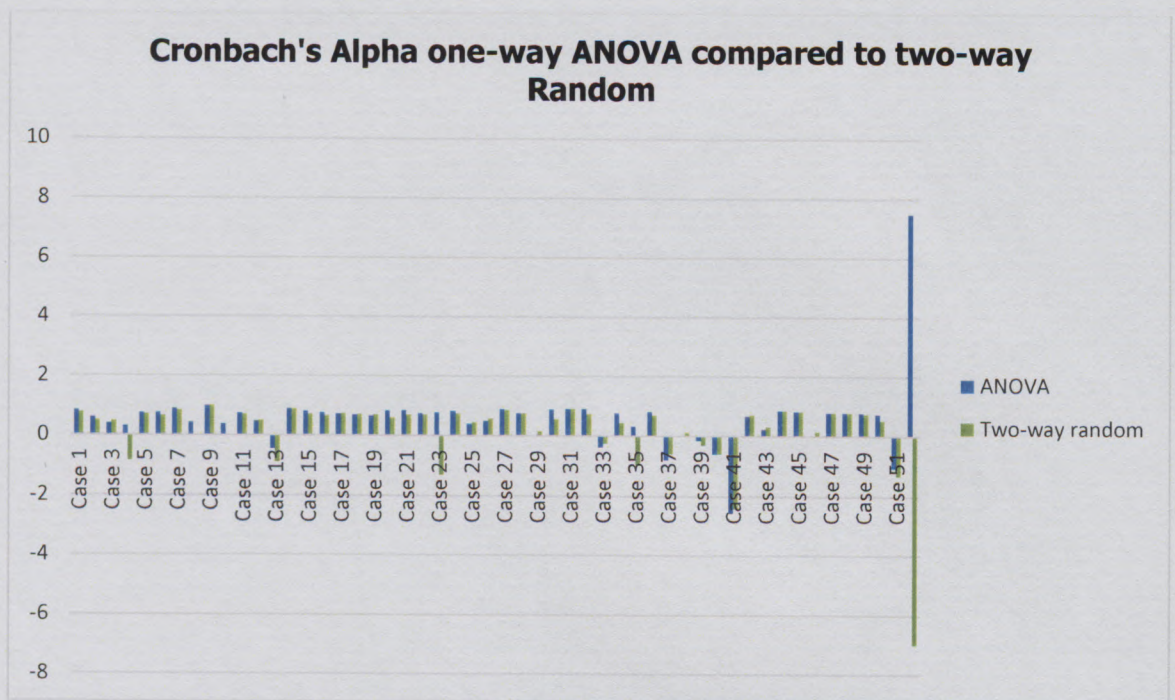


Figure 19- A comparison between one-way ANOVA and two-way random Methods when Calculating Cronbach's Alpha.

The researcher would like to point out that, although interesting, one should be wary of trying to extrapolate findings from the two-way model as there were only two raters and nine areas being tested each of the 52 times. This means that inconsistencies may be exaggerated.

That said, the two methods of calculating α produced similar results. This shows that there is good enough inter-rater reliability, as well as consistency to argue that the instrument needs refinement and does not need to be re-designed.

4.4.4 Summary of Part Four

During part four, the researcher presented the findings stemming from the CPAI mark sheets. Each of the 52 assessments completed using the UJ CPAI were captured on an Excel spreadsheet. Cronbach's Alpha was then

calculated for each assessment done. This means that Cronbach's Alpha was calculated 52 times. A good Cronbach's Alpha was found in the majority of the assessments. This shows good enough inter-rater reliability to continue with the refinement of the CPAI, rather than to redesign the instrument.

The researcher then calculated a simple correlation coefficient (Pearson's r) for each assessment area. Assessment area two was found to have a poor correlation and this was included in the agenda for the focus group discussion for further exploration. The researcher also found that the one-way ANOVA was comparable to the two way random Cronbach's alpha. This showed consistent results, whether reliable or unreliable.

4.5 Conclusion

During part one, the researcher described the format and structure of the UJ CPAI. The structure of the UJ CPAI follows nine assessment areas structured chronologically according to the components of a typical pre-hospital incident. These assessment areas were identified according to the components of a typical pre-hospital incident. The assessment areas of the UJ CPAI differ from other, similar assessment instruments used internationally. All of the assessment areas are well described in the context of emergency medical care and therefore it may be argued that their inclusion to the CPAI is logical.

The format of the UJ CPAI follows a narrative that informs a five-point rating scale for each assessment area. The narrative may be compared to the Dreyfuss' model of skills acquisition, as well as a logical description of what senior students at the Department EMC should be capable of when managing real-life patients. The five-point rating scale may also be

compared to the rating scale used in the Mini-CEX and, to a lesser extent, the instrument used by Tavares.

The researcher presented the results of the data analysis to test construct validity during part two. The results showed that the two-way random Cronbach's Alpha was low. This means that the ratings provided by the Subject Matter Experts were inconsistent. Also, in line with the poor consistency, it was found that, in some cases, participants awarded the exact opposite marks for the same assessment areas. This shows a high likelihood of poor construct validity. What this means is that it is likely that other tests, especially inter-rater reliability will also be sub-optimal. The researcher carefully looked at these findings and will discuss them when all of the data sets are corroborated.

The free responses that were generated from the open section of the questionnaire lead to the identification of a number of themes that were included in the agenda for the focus group discussion for further investigation.

In part three, the researcher presented another set of quantitative data. This data was generated by the completion of a self-designed, non-validated, evidence informed questionnaire. Seven subject matter experts agreed to complete the questionnaire. Eight of the nine assessment areas were awarded good ratings by at least five SMEs. Assessment area three was awarded a good score by only three experts. Assessment area eight was awarded the best ratings, and assessment area three was awarded the worst ratings.

During part four, the researcher presented the findings stemming from the CPAI mark sheets. Each of the 52 assessments completed using the UJ CPAI were captured on an Excel spreadsheet. Cronbach's Alpha was then

calculated for each assessment done. This means that Cronbach's Alpha was calculated 52 times. A good Cronbach's Alpha was found in the majority of the assessments. This shows good enough inter-rater reliability to continue with the refinement of the CPAI, rather than to redesign the instrument.

The researcher then calculated a simple correlation coefficient (Pearson's r) for each assessment area. Assessment area two was found to have a poor correlation and this was included in the agenda for the focus group discussion for further exploration. The researcher also found that the one-way ANOVA was comparable to the two way random Cronbach's alpha. This showed consistent results, whether reliable or unreliable.

In the next Chapter, the researcher will present the findings that emerged from the focus group discussion. A more detailed discussion of what these mean in the context of the study is provided in Chapter six.

CHAPTER 5- FOCUS GROUP DISCUSSION

5.1 INTRODUCTION

As mentioned in Chapter 3, a sequential model consisting of three different research processes were applied to investigate and describe the validity and reliability of the UJ CPAI. The first consisted of a literature review that provided background on the development and structure of the UJ CPAI. The second method made use of a questionnaire that provided quantitative data that was subsequently analysed to explore the content and construct validity for the UJ CPAI. The results of the abovementioned processes were presented in Chapters 2 and 4 respectively. This Chapter will focus on the third process, namely the focus group discussion. This chapter confines itself to a presentation of the findings that emerged from the focus group discussion, a more detailed discussion of what these mean in the context of the study is provided in Chapter six.

5.2 THE FOCUS GROUP DISCUSSION

The principle aim in all interviewing is obtaining valid and reliable information (Fielding, 2003). In this study a focus group discussion was conducted with academic staff members, clinical facilitators and external clinical assessors. The aim of the focus group was to further explore and describe participant's views and opinions on components of the UJ CPAI with regard to the construct and content validity, reliability and future use of the instrument within the department.

The focus group discussion was conducted at the Doornfontein Campus of the University of Johannesburg in the boardroom of the Faculty of Health Sciences. The focus group discussion was chaired by the researcher and lasted for 58 minutes. The recording was transcribed verbatim by the

researcher and yielded a transcription report of 27 pages. The transcript was analysed by means of simple content analysis as explained in Chapter three. The remainder of this Chapter will deal with the themes and trends that emerged.

5.3 MEASURES TO ENSURE VALIDITY RELIABILITY AND TRUSTWORTHINESS

The points identified for discussion (Annexure 11) were informed and carefully crafted to explore more deeply selected key findings that had emerged from the preceding research processes. As mentioned in Chapter 3, the focus group discussion was audio recorded and a verbatim transcript produced. Responses, themes and trends were then identified using thematic content analysis (Krueger, 2015). In addition the researcher conducted randomly spot checks comparing the audio-recording to the verbatim transcripts to further ensure the trustworthiness of the data (Shenton, 2004).

5.4 PARTICIPANTS IN THE FOCUS GROUP DISCUSSION

There were seven participants who were purposefully selected in an attempt to ensure that the important role players involved in the assessment of third and fourth year UJ Department EMC students were all represented. The main role players represented by these participants were:

- Three Academic staff members of the Department EMC not involved in the development of the UJ CPAI.
- Two Clinical Facilitators from the Department EMC

- Two Externally appointed clinical assessors

5.5 THE FOCUS OF DISCUSSION FOR THE GROUP

In Chapter one, the researcher stated that one of the questions central to the aim of this study was to investigate whether there is evidence of a need for adjustments and or refinements to the CPAI and if so what these were. In Chapter three, the researcher cited Creswell and Plano Klark and argued that, by applying a mixed methods approach to the research question the researcher is able to obtain a more complete analysis of the problem. In keeping with the argument, the researcher argues that a good approach of gathering valid evidence that will allow for making abovementioned recommendations would be a qualitative approach in the form of a focus group discussion.

Through analysis of the quantitative data set, the researcher identified certain key findings that warranted further investigation. These findings were related to the construct validity and the inter-rater reliability of the CPAI. The researcher used the focus group discussion to gather more evidence in order to i) identify those areas that needed to be adjusted and refined and ii) to recommend what type of adjustments needed to be made.

In addition to using the focus group discussion to investigate the findings of the quantitative processes further, this focus group discussion was also conducted to speak to the last objective of this study, which was to investigate the views and opinions of educators regarding the CPAI and its use for the assessment of clinical performance of emergency medical care students in the pre-hospital environment

A reference document (Annexure 11) was drafted which contained a number of points for discussion. The reference document would guide the focus group discussion toward the gathering of data applicable to the central aim and objectives of the study.

5.6 THEMES AND TRENDS THAT EMERGED FROM THE FOCUS GROUP DISCUSSION

The findings (in the form of themes) that emerged from the focus group discussion in relation to each of the four focus areas are presented below. A more detailed discussion of what these mean in the context of the study is provided in Chapter six.

FOCUS AREA ONE- CONSTRUCT VALIDITY OF THE UJ CPAI

As mentioned in Chapter 3, construct validity has been defined in this study as the ability of an assessment instrument to measure specific outcomes. This focus area was seen as important because scholars increasingly refer to the role that construct validity plays in other forms of validity as well as reliability (Bashook, 2005; Rosenthal & Westen, 2003).

In Chapter 4, the researcher cited Edens and Swoboda and stated that a very important step in the development of any workplace-based assessment instrument is to ensure that it is designed well. In order to do this, it was necessary to determine exactly which criteria the instrument will assess, as well as how the instrument will assess these criteria. The argument was made that if the construct validity of an instrument is poor, it may cause assessors to draw incorrect or inconsistent conclusions and adversely affect validity and reliability.

If recommendations for the refinement are to be made responsibly, in a manner that will improve construct validity, the views and opinions of the users of the instrument should be included.

Background to the inclusion of the first discussion point

When the data of the questionnaire was analysed, one of the pertinent themes in the open response section was that Assessment Area one (Preparation) was only really applicable for the first case (4). This is consistent with the researcher's comment on the same assessment area earlier in Chapter 4. This first question was therefore formulated in order to i) investigate the opinions of the users of the CPAI in order to see whether the feelings are consistent with the quantitative results and ii) how to appropriately refine the narrative of the assessment area.

The researcher first provided some background to the question. An opening statement was made that, during completion of the questionnaire, there was a strong feeling that assessment area 1 is only applicable for the first case. The researcher proposed that this problem may be corrected by doing a number of things. These include that the mark awarded for this area on case one be carried over, or that the narrative be changed so that it can be used on all of the cases being assessed. In order to decide what to do, it may be valuable to ascertain whether there are any pertinent areas the participants can identify that are only applicable to the first case and then decide on follow-up questions.

<p>Discussion Point 1 - Are there specific areas in this narrative that could be singled out that is only applicable to the first case?</p>
--

Two main themes stemmed from this discussion:

The first theme was that **the narrative that informs the assessment of preparation (assessment area one) is only applicable to the first case.**

"... if you were to single out... say, for example, the vehicle's tyres are polished... they are not going to polish [them] between every call..." – Clinical facilitator from the Department EMC.

One participant argued for this assessment area should be removed from the instrument altogether, this was not well supported.

"...my view... potentially is that this area is *entirely* excluded out of a clinical practice assessment instrument." – Academic staff member from the Department EMC.

However the remaining participants felt that the area should stay but the narrative should be adjusted so that it can be used for all incidents. This included the removal of criteria such as polishing the tyres of the vehicle and cleaning the vehicle, and explicitly identifying preparation or the performance of clinical duties.

"... if your boots aren't polished, it *actually* doesn't impact... on [clinical performance]... But, if you went on a call, used equipment, forgot about it, go to the next call, you literally *can't* do that clinical intervention..." - Academic staff member from the Department EMC.

"Maybe the focus could be on preparation that has a direct link to clinical performance." - Academic staff member from the Department EMC.

The second theme that emerged was that **assessment area 2 (that deals with professional presentation) is not seen as applicable to clinical performance and therefore should not form part of the CPAI.**

There was a general feeling that, although professional presentation is important, it did not directly influence clinical performance. There was a feeling that professional presentation should perhaps form part of a different feedback and assessment strategy.

"I'm not sure whether what you look like and whether your equipment is packed before every call contributes to your clinical competence... It's important – I'm not taking that away, but I don't know if that must maybe form part of something else..." - Academic staff member from the Department EMC.

"...But I don't think that that has anything to do with your clinical competence..." - External clinical assessor

"... as a person that used the tool, I've found the first two to be silly... not silly... um important but..." - Academic staff member from the Department EMC.

"...because I can tell you now that... as an educator, I can address somebody's uniform, the way he speaks, in the case of one day – 'ok so uniform is terrible – tomorrow it will be sorted out..." - Academic staff member from the Department EMC.

Background to the inclusion of the second discussion point

An important trend that emerged from the free response section of the questionnaire was that the rating scale used in the UJ CPAI was confusing. Responses in the questionnaire indicated that it was easy to understand the narrative on when to award marks 1-4 (does not comply until exceeds expectations). It was unclear how students should perform to be awarded a mark of 5. The reason this theme is important for this study and warrants further investigation is because the researcher argues in Chapter 4 that the assessors who will be using the assessment instruments need to clearly understand how to award marks for students. The rating categories should be written in such a way that is clear and easy to understand. That way, not only does the construct validity increase, but the likelihood of improved inter-rater reliability also increases.

Discussions point two - Do you think that the instrument should include five rating categories instead of four?

Again, two themes emerged from this question. One theme was that the participants felt that five rating categories are not necessary. The other theme was that there is general support by the participants for three rating categories.

Theme: The participants felt that five rating categories are not necessary

The feeling by the participants was that five rating categories increased the confusion. There was a strong feeling that the fifth rating category was too vague, and therefore it was difficult to identify when students were eligible to be awarded a rating of five.

"I don't think you need a five."- Academic staff member at the Department EMC.

"For me, in every case I found that... what I would say... you know... what the *top end* was for everyone was the narrative for level four..." - Clinical facilitator from the Department EMC.

Theme: There is general support by the participants for three rating categories

There was a feeling amongst the participants that, with less rating categories to choose from, they would find it easier to decide which marks to award students.

"... I think when you have fewer categories, you can say with more certainty..." - Academic staff member from the Department EMC.

"the only thing it will probably come back to is the way to was set up initially is that it is asymmetrical, so meet minimum standard is... if you take the five out, there are two below the three, but only one above the three..." - Academic staff member from the Department EMC.

"you can take the two out and just have a few... not yet competent, competent, best practice" - Academic staff member from the Department EMC.

"... a one, a two and three. Where a one will be... be not yet competent, or needs improvement, a two will be competent and a three will be you exceeded expectations..." - Clinical facilitator from the Department EMC

One of the comments that the researcher found interesting, but that was not in the scope of the focus area was a comment made regarding the descriptions of the levels of performance.

Sorry - from the student's perspective – if I put a student hat on – what does needs improvement mean? - Academic staff member from the Department EMC.

"...even if you are operating at a four, or five, you still need improvement...." - Academic staff member from the Department EMC.

FOCUS AREA TWO- CONTENT VALIDITY OF THE UJ CPAI

Content validity refers to the question "does the assessment instrument adequately assess what it was designed to assess?" As mentioned in Chapter 4, content validity is considered by some authors to be an important form of validity for assessment instruments (**cf. Chapter 4**). Other types of validity may influence content validity, but none can replace it.

Background to the inclusion of discussion point three

When analysing the quantitative data from the questionnaire, only one assessment area was rated as having good content validity by all of the participants. This is consistent with the literature in that content validity cannot be proven once off, but should rather be improved constantly.

That said, assessment outcome three (History taking) was considered by three of the seven respondents (43%) to have good content validity. Question two was developed because the above indicated to the researcher that it is necessary to clarify why the SMEs thought that

content validity was poor for assessment outcome three, as well as how content validity could be improved.

Discussion point three started with an introduction, where the researcher indicated to the participants that assessment outcome dealing with history taking was considered by four out of the seven (57%) of the participants to have poor content validity. The feedback provided in the questionnaire, the participants felt that they got the impression that two outcomes were assessed with this assessment outcome. An example of this is that there is a difference between communicating well, and obtaining a good history. The narrative therefore did not address the outcome of history taking adequately.

Discussion point three - Do you feel that the content validity of assessment area three (history taking) was poor because two outcomes were assessed?

Two themes emerged from this question. **The one theme was that participants were of the opinion that the way the narrative for the area that assessed the students' ability to obtain a good history was structured made it difficult to identify exactly what the student is being assessed on.** The other theme was that **participants made the argument that some of the constructs being assessed should be moved to form part of other assessment areas.**

"my feeling is that the minute you... uh... you... increase the number of verbs in the assessment, the more complex it becomes" - Academic staff member from the Department EMC.

you actually asking me in assessment area three to assess three very different...um... ideas or.... sorry... actions that these guys are going to be performing” - Academic staff member from the Department EMC.

“history taking is actually a quite an important *independent* entity that needs to be assessed” - Academic staff member from the Department EMC.

“I agree with [previous comment]” – External Clinical Assessor

“...did you get what you needed and are acting on it? Versus your... professionalism.” – External Clinical Assessor

“I know when we did this simulation task team discussion, history taking can be put under a different area... form part of diagnostics... communication and general patient interaction should fall under professionalism... as opposed to under actual history taking.”- - Academic staff member from the Department EMC.

The researcher did note, however, that at times the participants were unclear about what each assessment area entailed. The comment on the statement is that history taking and physical assessment do not form part of the same assessment area or narrative. They are separate entities on the UJ CPAI.

“... the descriptions in the rubric are packaged together, so the guy who is poor, or does not comply, it assumes he is poor at everything... like you are saying... you take a good history, but does a poor physical assessment, you then can't fit into one of these categories, all poor, all ok...” - Academic staff member from the Department EMC.

"I think... although you don't want to go crazy with assessment areas, but maybe history and physical assessment should be... separate... they're too complicated." - Academic staff member from the Department EMC.

FOCUS AREA THREE- INTER-RATER RELIABILITY OF THE UJ CPAI

Background to inclusion of discussion point four

The reason discussion point four was included in the agenda was because the results from the completed UJ CPAI mark sheets showed good inter-rater reliability in only 67% of the cases. The researcher identified the need to investigate possible reasons for the sub-optimal number of assessments that showed good inter-rater reliability. Some authors argue that some common reasons for poor correlation coefficients during inter-rater reliability assessments include poor or incomplete training of assessors on the use of the assessment instruments. This caused different assessors to award marks differently for the same assessments.

Discussion four started with an introduction. The researcher firstly explained shortly what inter-rater reliability is and that inter-rater reliability is commonly measured using a correlation coefficient. This means that the scores awarded to a student across all nine assessment outcomes by two assessors were compared and a correlation coefficient was calculated. The researcher then informed the participants that a correlation coefficient that is considered to show good correlation or agreement between the assessors is 0.6 or more, and that, out of the 52 cases documented, a good correlation was seen 34 times. This means that the data showed that the assessors awarded the same marks in less than 75% of the cases the CPAI was used.

The researcher then stated that authors are of the opinion that one reason for sub-optimal inter-rater reliability findings between assessors include that the assessors were not adequately trained on the use of the assessment instrument.

Discussion point four - Did you find that inadequate training might have been the reason that less than 75% of the assessments had good inter-rater reliability?

Three themes emerged from this question. The first theme was that the participants felt that training did not impact the inter-rater reliability of the instrument. There was a general feeling that the reason for the poor inter-rater reliability was the way in which the narrative was written. The participant felt that they perceived the assessors who used the CPAI to assess students to be set in their ways.

Theme: The participants argued that training did not impact the inter-rater reliability of the instrument.

There was a feeling that incomplete training was not the reason for the cases with poor inter-rater reliability. Some of the participants felt that training alone will not remove assessor bias.

"there's other stuff about training for simulations, which shows it *doesn't* work..." - Academic staff member from the Department EMC.

"...ja, ous (sic) are random, they get trained and they *still* give their own marks." - Academic staff member from the Department EMC.

"My feeling is... I might differ with these guys that say you need training..." - Academic staff member from the Department EMC.

Theme : A proposed reason for the poor inter-rater reliability was the way in which the narrative was written.

This theme repeated again. A theme with question one was also that the way in which the narrative was written was confusing and lead to different assessors awarding marks differently. Some of the participants referred to question one again as well.

"I think this thing spoke to those categories..." - Academic staff member from the Department EMC.

"...it's not helped by the fact that there are different constructs in a category..."- Academic staff member from the Department EMC.

"You see... my view on the categories as well is you can have quite a naive assessor... when he looks at the descriptors, he goes: "student appears"... I'm just using this as an example... "student appears anxious, and shows a lack of confidence when trying to interact with the patient." he ticks it... [pause]... then later on, "their bedside manner and body language needs work" it's in a different category, but he ticks it... sometimes the assessor does not know... now... where does the majority of bad or good stuff lie, so which category do I need to... need to go for..."- Academic staff member from the Department EMC.

Theme: Assessors are perceived by the participants to be set in their ways.

There was a general feeling amongst the participants that each narrative still needs to be interpreted. Also, in some cases, assessors will award

marks according to their own set of beliefs and opinions regardless of what the narrative advises.

"... you can take a rigid paragraph and we'll all interpret it differently..." - Academic staff member from the Department EMC.

"Specifically one thing that stood out for me was... assessor was very set on in order to score competently for... uhhh clinical assessment, the... the student needed to take a pulse on the foot of the... you know, leg with a tibia fracture and therefore it scored a one or whatever... you know, whereas everything else had been done quite well, so I think that sort of outlying..." – Clinical Facilitator at the Department EMC

Background to the inclusion of discussion point five

The introduction for discussion point five was that an example of an area that was associated with a very different correlation coefficient was assessment the area of professional presentation (Assessment area two). The above assessment area had a correlation coefficient of -0.37. This means that, one can say with a small degree of certainty that the marks awarded by the raters is likely to differ. This was also the only area with a coefficient of less than 0.4. One argument may be that the assessors understood the narrative very differently and therefore awarded marks very differently.

<p>Discussion point five - Why do you propose assessment outcome 2 (professional presentation) had such a low correlation coefficient?</p>

Again, the overarching theme was that **the assessors have differing understandings about what the narrative means.**

"I'm not surprised, to be honest. If you... if you wanna example of... um... what I was talking about earlier about set in their ways probably is this..."

- Academic staff member from the Department EMC.

"Ja... and I tend to agree... this... this also speaks to... the blurring of categories again." - Academic staff member from the Department EMC.

"I can be telling you... with... well... absolute certainty... there'd be two students – one student would walk into a patient's... uh... room and say: 'good morning sir, we are from the emergency services'. Then another student who sits on the bed – 'howzit' – he taps the guy on the shoulder... and assessor – although both are completely appropriate approaches because there's one guy – that goes the 'good morning sir', until 'you can call me by my first name' and all that type of stuff... that is where I *think* you did see this difference... and when *in fact*... if you look at it, these guys weren't *that* different..." - Academic staff member from the Department EMC.

FOCUS AREA FOUR- SUPPORT FOR THE FUTURE USE OF AN
ASSESSMENT INSTRUMENT SUCH AS THE UJ CPAI

The reason focus area four was included in the focus group discussion was to speak to one of the objectives of the study, which was to report on the opinions of the academic staff members, clinical facilitators and external examiners at the UJ Department Emergency Medical Care regarding the continued use an instrument such as the UJ CPAI as an assessment instrument.

Discussion point six - Do you feel that the use of this kind of instrument should be continued in the programme?

Three themes emerged from this question. The first theme was that there was a strong feeling that there is a need to implement an instrument such as the UJ CPAI. The second theme was that the participants felt that the CPAI needed to be improved before it could be implemented. There was also an overwhelming response by the participants that the time and effort required to complete the document needed to be decreased

Theme: There is a need for an instrument such as the UJ CPAI.

All of the participants agreed that there is a need for an assessment instrument such as the UJ CPAI. The use of the UJ CPAI allowed not only for assessment, but also for good feedback. This in turn, could guide reflection and learning.

"...it's important especially for the students as [previous participant] also mentioned is that at least now they've got something to go back and reflect on." – Clinical Facilitator at the Department EMC

"...so I think with this tool, even... even if that feedback is negative, they'll always go back and say "I've scored negatively on that point... why did I score negatively..." – Clinical Facilitator at the Department EMC

Theme: The CPAI needed to be improved before it could be implemented.

There was agreement that the UJ CPAI needed to be improved before it could be implemented.

"I think uh... if it's simplified – reduce the number of categories, separating those issues...." - Academic staff member from the Department EMC.

Theme: The time and effort required completing the document needed to be decreased

A strong feeling was that procedure to complete the UJ CPAI after each assessment needed to be streamlined. The time and effort required completing the document needed to be decreased.

"... I find this very cumbersome... um... it's... it's not a user friendly thing to fill out..." – External Clinical Assessor

"I must... I actually support that... consider the environment we're working in... There's nothing worse than, on a clipboard, turn the pages over..." - Academic staff member from the Department EMC.

"if you leave the criteria there, i know that all i have to do on call one is it's a one there, a two there...quickly quickly done, get all the numbers, put in and it's done." – External Clinical Assessor

5.7 CONCLUSION

Some of the important themes which emerged from the focus group discussion centred on the way that the narrative of the UJ CPAI was written. There was general support for the use of an assessment instrument such as the UJ CPAI. The instrument was seen as useful not only for assessment, but also for feedback. That said, there was also a strong theme that the UJ CPAI in its current format is cumbersome to complete and the process needed to be streamlined before

implementation. There was general agreement that some of the assessment areas of the UJ CPAI were only suited to assess the first case. Also, there was a feeling amongst the participants that the narrative was written in such a way that it confused assessors at times, as well as caused assessors to reach different conclusions when assessing the same student. The assessors also felt that five rating categories were too many. There was general agreement that three rating categories would be more appropriate for the UJ CPAI. In other words, the participants felt that the UJ CPAI would be a valuable assessment instrument for use in the Department EMC, provided that a number of identified changes and refinements are made to the CPAI.

5.8 SUMMARY

As mentioned in Chapter 3, a sequential model consisting of three different processes were applied in order to gather data for this study. This Chapter presented on the qualitative data collected and analysed for this study. The quantitative data showed that the UJ CPAI was not capable of producing valid and reliable results consistently and therefore warranted to be adjusted and refined. The researcher needed to gather more evidence in order to i) identify areas that needed to be adjusted and refined and ii) to recommend exactly which adjustments need to be made. A focus group agenda was therefore carefully constructed using pertinent findings from the quantitative research processes that needed to be explored further. Nine assessors who used the assessment instrument to assess the UJ BHS EMC students were invited to take part in the focus group discussion. Of the nine invited, eight were able to participate. The focus group discussion took place in the seventh floor boardroom of the John Orr building at the UJ Doornfontein Campus. The focus group discussion was audio recorded, and transcribed verbatim. The focus group discussion lasted 58 minutes and produced a transcript of 27 pages. The

transcript was then coded using simple content analysis. Themes and trends were then identified and discussed.

In the next Chapter, the researcher will complete the final step in the research process, which is to combine the data gathered through the qualitative as well as quantitative methods. This will be done by triangulating of the data and corroboration of the data sets. The pertinent findings will then be integrated and recommendations regarding the structure, process, outcomes and future use of the CPAI within the UJ BHS.EMC Degree programme will be made.

CHAPTER 6 –DISCUSSION OF CENTRAL FINDINGS

6.1 INTRODUCTION

The aim of this study was to investigate the validity and reliability of the UJ CPAI. This was done by through the application of a prospective, descriptive, sequential mixed method design.

Firstly, a literature review was completed in order to frame the study within the context of existing knowledge. Theories that guide assessment, especially the assessment of clinical performance, were discussed and used to put forward the argument for the need to validate an instrument for the assessment of clinical performance in the South African pre-hospital emergency care education domain. This literature was presented in Chapter 2. After the literature review, the structure and format of the UJ CPAI was described by comparing the assessment areas of the UJ CPAI to other internationally validated assessment instruments.

The researcher then invited seven participants (four academic staff members from the UJ Department EMC, two clinical facilitators at the UJ department EMC and one external clinical assessor) to complete a self-designed, evidence informed questionnaire (Annexure 6). The survey questionnaire focused on exploration of the construct and content validity of the UJ CPAI. Whilst many different types of validity have been identified and explored in the literature review, the survey questionnaire focused on construct and content validity as these two types of validity have been seen as overarching themes in validity. Although other validity themes exist, it is argued that none can replace content and construct validity (Salkind, 2004). The results emerging from analysis of the responses to the questionnaires were presented in Chapter 4.

Following the on from exploration of the validity of the UJ CPAI through the survey questionnaire, the researcher proceeded to investigate the reliability of the instrument. Inter-rater reliability of the UJ CPAI was tested through its introduction onto the UJ Departmental Primary Response Vehicle during clinical learning shifts from February 2015 until November 2015. A total of 104 completed CPAIs representing 52 incidents were retrospectively analysed and a Cronbach's Alpha was calculated. The results of this analysis are presented in Chapter 4. This was then followed by a focus group discussion that sought to investigate and describe the views and opinions of educators regarding the CPAI and its use for the assessment of clinical performance of emergency medical care students in the pre-hospital environment. The agenda for the focus group discussion was carefully and purposefully designed being informed by results and themes that had emerged following the literature review and analysis of the questionnaires and completed CPAIs. The focus group discussion and the results thereof were presented in Chapter 5.

In this penultimate Chapter the researcher attempts to triangulate the data and findings of the three research processes summarised above allowing for identification and discussion of the central findings. Triangulation allowed the researcher to be more confident about the results and findings from this study (Cassim, 2014). Through triangulation answers to the main research questions and central findings are able to be unpacked, described and discussed in this chapter and the sections that follow.

6.2 VALIDITY OF THE UJ CPAI

As mentioned Chapter One, the first objective of this study was to comment on the validity of the UJ CPAI. Two types of validity were investigated- construct and content validity.

6.2.1 Construct validity of the UJ CPAI

As mentioned in Chapter 4, Construct validity is seen as the overarching theme of validity, and refers to the degree to which an assessment instrument is designed to measure specific areas (Rosenthal & Westen, 2003).

The researcher invited nine prospective participants to complete a questionnaire regarding the validity of the UJ CPAI (Annexure 6). As explained in Chapter 3, the questionnaire in this study required participants to rate the extent or depth to which they felt that the narrative of each of the nine assessment areas making up the UJ CPAI allowed them accurately score the student's performance. From the responses provided by the participants, the researcher calculated Cronbach's Alpha (α) using a two-way random model to quantify the consistency of the responses provided by the participants expressed as a numerical value (Norusis, 2003; Rosenthal & Westen, 2003; Landis & Koch, 1977; Salkind, 2004). The researcher was then in a better position to comment on the construct validity of the UJ CPAI.

When the questionnaire was analysed, the two-way random Cronbach's Alpha was low. Consistent with the low α , the researcher found that the participants awarded ratings inconsistently. An example of this is that one participant awarded a Likert-type rating of five to Assessment Area three (Communication, History Taking and general patient interaction), meaning that the participant felt that the narrative perfectly allowed them to accurately score the student's performance. For the same assessment area, another participant awarded a rating of two, meaning that the narrative poorly enabled them to assess the student's performance. Also, participants four and seven awarded almost exactly opposite ratings throughout the questionnaire. The absence of a pattern to the responses

means that the low Chronbach's Alpha is accurate. Although other reasons, such as the small sample size could have also contributed to the absence of a pattern, however another likely reason for a low α may be poor construct validity.

These results seem to support the idea that the participants interpreted the narrative of the UJ CPAI differently. The researcher argues that the narrative does not provide sufficient explanation of exactly what is being assessed. In further support of this finding was found in the free response sections of the questionnaire where participants indicated that the CPAI narrative for some assessment areas was too vague to elicit a deeper response.

What the researcher also found was that some of the participants noted that they found the mark allocation confusing. The participants noted that there was very little congruence between the narrative informing a mark of one and the narrative informing a mark of five. Participants struggled to find the links between the narrative describing the best performance, and the narrative describing the worst.

The low correlation coefficient showing poor consistency, the absence of patterns in the awarding of marks and the comments regarding the mark allocation lead the researcher to argue that there were areas of the UJ CPAI that had poor construct validity (Edens and Swoboda, 2013, Rosenthal & Westen, 2003). On further investigation, four areas were identified where the format of the UJ CPAI decreased construct validity. These four areas were: i) the narrative of assessment area one (preparation), ii) the narrative for assessment area two (professional presentation), iii) the narrative for assessment area three (history taking, communication and general patient interaction) and, iv) the five-point scale used to award marks to the students.

These four identified areas will be discussed in more detail in the next part of the chapter. The researcher will put forward possible reasons why each of the abovementioned areas decreased the construct validity of the UJ CPAI and proposed ways to possibly improve these four areas.

i. Assessment area one – Preparation

Preparation for the task at hand is arguably one of the most important duties the clinician may perform on a call-to-call basis (Sanders, 2001). This means that it is important to assess a student's ability to adequately prepare for an incident when assessing clinical performance. That said, during the focus group discussion, there was consensus by the participants that the performance criteria contained in the narrative only allowed the assessor to assess the student for the first assessment of the day. The general feeling was that the assessment criteria in the narrative was not relevant for subsequent assessments.

Some of the performance criteria included in the narrative were punctuality for shift, the cleanliness of the vehicle, and required levels of disposable medical stock. While these are important for the start of the shift and for the first assessment, the researcher agrees with the focus group discussion panel that these criteria would only be valid once. For subsequent assessments, the narrative provided little guidance and it is understandable that assessors found it difficult to award marks during subsequent assessments. Because of the lack of guidance, assessors did what they thought best for subsequent assessments. This lack of guidance decreased construct validity of the instrument.

ii. Assessment area two – Professional presentation

During the focus group interview, there panel members were of the opinion that, while professional presentation is important, it may not be relevant to clinical performance. These views are not well supported by literature. Roberts (2011) argues that professionalism is, in fact, an important component of clinical performance. Professionalism is also an assessment area of the mini-Clinical Evaluation Exercise (mini-CEX) (ABIM, 1995). Although it is posed that professionalism is considered difficult to pin down (Roberts, 2011), authors include entities such as compassion, respect, integrity and excellence to describe professionalism (Roberts; 2011; HPCSA, 2008). Faure (2014) also cites the HPCSA when stating that professionalism refers to communication and behaviour in addition to physical appearance. Not only has it been shown that professional appearance may lead to a greater degree of trust the patient has in the health care provider, it is a necessity in terms of the South African health care setting (Faure, 2014, Sanders et al., 2001).

The narrative for this assessment area may have been written in a manner that a limited range of criteria is assessed with this assessment area. Criteria that was included to assess area two (professional presentation) were entirely dedicated to physical presentation such as completeness and appropriateness of uniform, hair, nails and shoes. When taken into consideration that professionalism encompasses more than just physical appearance (Roberts; 2011; HPCSA, 2008, Faure, 2014) the researcher poses that the assessment area could be changed to "professionalism" and more applicable criteria should be used to adequately assess the area. This argument ties in with the next assessment area– assessment area three.

iii. Assessment area three – History taking, communication and general patient interaction.

When asked about this in the focus group discussion, the participants were of the opinion that the way the narrative for this assessment area was structured made it difficult to identify exactly what the student is being assessed on. As mentioned earlier, this supports the argument made earlier that assessment criteria or “behaviours” need to be well defined and relevant to the assessment area in order for assessors to be able to use the instrument properly (Beckman & Cook, 2005; Edens & Swoboda, 2013). Behaviours used to assess assessment areas need to be directly related to the assessment area (Beckman & Cook; 2005, Guyatt et al., 1993). It is also demonstrated that the assessment criteria should be simplified and focused to assess a specific area. When too many criteria are included, or when criteria become blurred, the instrument will no longer be capable of producing valid results (PMETB Workplace based assessment Subcommittee, 2005).

In this case, the assessors felt that the assessment area contained too many criteria that the students were assessed on. An example of this is that history taking and communication are two very different things. Communication refers to things such as professionalism, levels of comfort and anxiety, body language and tone of voice when engaged with history taking. History taking refers to gathering information regarding the history of the presenting illness (Bickley, 2003). The argument made by the participants was that one could obtain a good history, but communicate poorly and vice versa. There was agreement during the focus group discussion that the part of the narrative dealing with communication and general patient interaction is, in-fact, part of professionalism.

The health history is considered one of three critical areas of clinical proficiency (Bickley, 2003). The health history, including the gathering information regarding the history of the presenting illness, forms the cornerstone of putting findings into context, making diagnoses and formulating management strategies (Bickley, 2003). In fact, the ability to obtain a comprehensive health history and communicate with patients effectively is one of only two assessment areas that were included in all four assessed workplace-based assessment tools (Tavares et al., 2013; ABIM, 1995; Mitchell et al., 2011; Mamelok, 2009). For this reason, the researcher posits that the criteria of the assessment area should be narrowed to focus only on the students' ability to obtain a comprehensive health history. The criteria used to assess communication may potentially be better suited for inclusion in another assessment area.

The ability to communicate well may be better suited to professionalism, as put forward by the focus group discussion. Roberts (2011) agree with this, stating that professionalism encompasses entities such as compassion, respect, integrity and excellence (Roberts, 2011). The HPCSA and patients' rights charter refer to clinicians showing a positive disposition towards patients and those health care providers should demonstrate courtesy, human dignity, patience, empathy and tolerance (HPCSA, 2008). The researcher therefore argues that it would make sense to include effective communication as a criteria used to assess professionalism, as it is more applicable there.

iv. Rating scale of the UJ CPAI

One of the important steps necessary for the development of good assessment instruments is the development of an appropriate rating scale (Edens and Swoboda, 2013). The UJ CPAI followed a five-point rating scale. Clinical assessors observe the student as they manage a patient in

the prehospital environment and thereafter rate their performance in each of the nine areas using a five-point rating scale. Assessors award marks from 1 to 5 with 1 being poor and 5 the best possible performance. Assessor's allocations are guided by narrative that describes the typical actions that link best to each of the possible scores.

When the researcher proposed possible causes for the poor construct validity and sub-optimal inter-rater reliability of the CPAI during the focus group discussion, it was noted that the discussion panel members found the five-point scale used in the UJ CPAI too vague. There was a general feeling that a three point scale would be better suited for the UJ CPAI.

Different interpretations of the performance criteria contained in the narrative of the UJ CPAI could have caused assessors to award different marks for the same performance of a given student, as described by Bashook. Bashook (2005) states that assessors award marks differently when asked to rate how well students performed certain tasks. Well defined guides explaining how marks should be awarded are considered to be the most important task when developing workplace based assessment instruments (Edens and Swoboda; 2013). Simply put, when the mark allocation is clear and the room for misinterpretation is minimised, assessors will award marks more consistently. This will indicate improved construct validity.

The researcher therefore agrees with the focus group participants that a simpler, more concise three-point rating scale may lead to improved construct validity, as well as reliability of the UJ CPAI.

6.2.2 Content validity of the UJ CPAI

Content validity refers to how well an assessment instrument is able to assess the different assessment areas. A number of different methodologies to determine content validity have been described. A common method used is to make use of Subject Matter Experts (Sireci & Faulkner-Bond, 2014). As mentioned earlier in this Chapter, the researcher invited seven participants to complete a questionnaire pertaining to the validity of the UJ CPAI. The researcher calculated the internal consistency between responses to the survey questionnaire in order to comment on content validity. The one-way ANOVA method used to calculate Cronbach's alpha described the degree to which the participants awarded the same responses (Norusis, 2003).

In addition to calculating Cronbach's Alpha, the researcher also described the number of times SMEs awarded good Likert ratings for assessment areas. When five out of the seven (71%) of the participants awarded a good rating (four or five), the researcher argued that the content validity was high for that assessment area.

The questionnaire included seven participants to award Likert-type ratings for nine assessment areas. This means that each participant was required to provide nine responses. The total number of Likert-type responses was 63 (seven participants awarding nine ratings each). Of the 63 ratings awarded, a good rating of four or five was awarded 78% of the time (n=49). The one-way ANOVA Cronbach's Alpha was 0.752 (CI 95%, upper bound 0.911, lower bound 0.383). This suggests that, overall, the participants agreed that the UJ CPAI had good content validity.

The assessment area with the poorest content validity was assessment area three (History taking, communication and general patient

interaction). Less than half of the participants awarded a good rating for this assessment area. This makes sense, as assessment area three was described as having poor construct validity as well. When discussing the construct validity of assessment area three, the researcher indicated that the focus group participants were of the opinion that the way the narrative for this assessment area was structured made it difficult to identify exactly what the student is being assessed on. It would therefore be expected that the participants awarded poor ratings for the extent to which the narrative for assessment area three enabled them to assess area.

The assessment area that was awarded the most ratings of four or more was assessment area eight. All of the participants awarded a good rating for this area. Interestingly, assessment area eight produced the second highest Pearson's correlation coefficient. This is not surprising and the significance of this will be made clear in the next section that discusses the findings relating to reliability of the UJ CPAI.

6.3 RELIABILITY

Reliability refers to the ability of an instrument to deliver the same results repeatedly. Different forms of reliability exist. For this study, the researcher investigated inter-rater reliability. Inter-rater reliability refers to the degree to which different assessors (two or more) agree when awarding marks to a student using the same instrument (Viera & Garrett).

Inter-rater reliability was tested by comparing the marks awarded to students by two clinical assessors using the CPAI. As mentioned in Chapter three, the CPAI was used on the UJ Departmental Primary Response Vehicle from February 2015 until November 2015. Each of the students working a clinical learning shift on the UJ PRV was assessed

independently by two clinical assessors. This means that each assessor used the CPAI to assess the clinical performance of students. Each individual assessor was blinded to the marks awarded to the students by other assessor. The researcher collected all of the mark sheets after the completion of each shift and captured the data in a Microsoft Excel® spreadsheet. This spreadsheet was shared with a statistician in order to calculate Cronbach's Alpha (α) for each completed assessment, using one way ANOVA. A Chronbach's Alpha was calculated for each of the 52 cases that were completed. An α of 0.60 and a 95% confidence interval was used for all 52 cases. The researcher also calculated a Pearson's correlation coefficient for each of the nine assessment areas. The rationale for this is explained in Chapter three.

The results from the 52 analysed cases showed that an α of more than 0.60 was found in 67% of the cases it was used to assess. Although this means that there was good inter-rater reliability for more than half of the assessments done, 67% was less than the identified target of 80% needed to show good enough inter-rater reliability. This means that the researcher was unable to demonstrate that assessors using the UJ CPAI was capable of producing sufficiently reliable results for it to be formally introduced as an assessment instrument for the BEMC programme. In addition to that, only three out of the nine assessment areas showed a correlation coefficient of 0.600 or more. This strengthens the argument that the inter-rater reliability between the assessors using the UJ CPAI was low.

Three reasons are put forward that could have caused inter-rater reliability to be low. These were i) low construct validity, ii) Inadequate assessor training on the use of the UJ CPAI and iii) assessor bias when awarding marks to students. The effects of these three possible reasons are briefly discussed below.

6.3.1 Construct validity

Poor construct validity may cause poor inter-rater reliability (Bashook, 2005; Edens and Swoboda, 2013; Rosenthal & Westen, 2003). In the case of the UJ CPAI, the finding was no different. The areas that were associated with poor construct validity was associated with poor inter-rater reliability as well. During the focus group discussion, on the topic of about inter-rater reliability, the participants frequently referred to the construct validity of the CPAI as a cause for the low inter-rater reliability. Conversely, assessment area eight that showed good construct validity also showed good inter-rater reliability. The researcher therefore cites Rosenthal & Westen, Edens and Swoboda, Beckman and Cook and Guyen in arguing that, should the format of the UJ CPAI be refined in a manner which will improve construct validity, inter-rater reliability will also improve.

6.3.2 Inadequate Assessor training

There was a general feeling amongst the focus group participants that training assessors on how to use the UJ CPAI would not improve inter-rater reliability. The focus group discussion members stated that that training does not improve compliance with assessment narrative.

The researcher argues differently. The researcher agrees with the authors that showed that adequate training works well in eliminating assessor bias (Bashook; 2005, PMETB Workplace based assessment Subcommittee; 2005, Tavares et al. 2013). With proper training, assessors can understand why it is important to follow the narrative the way that it was written. The developers of the instrument (or, in the case of this study, the researcher) could also explain exactly what is assessed under each

assessment area, so that assessors could use the assessment instrument to award marks instead of relying on 'gut feel'. The researcher argues that, not only will assessor training improve reliability of the CPAI, but the credibility of the instrument will also improve (Prober, 2011; SAQA, 2015).

During the focus group interview, it became evident to the researcher that at times the participants were unclear about what each assessment area entailed. An example is that participants spoke about history taking and physical assessment as one assessment area. On the UJ CPAI, history taking is in-fact different from physical assessment already (Annexure 3). With more elaborate training, the researcher argues that assessor confusion regarding what exactly is assessed under each assessment area would have been improved.

6.3.3 Assessor bias

The participants of the focus group discussion mentioned that assessors in some cases awarded marks differently to the narrative when the student omitted a procedure they thought was very important. The participant referred to an instance where a student assessed the fractured leg of a patient. The narrative indicated that the student completed the majority of the necessary investigations and therefore should be awarded a mark of three for that assessment area. The assessor, however, awarded a mark of one since the student omitted a single assessment that the assessor felt was very necessary. In this instance, one assessor would have awarded a mark of one and the other a mark of three, this large difference between marks awarded by the two assessors decreases the inter-rater reliability, as well as the credibility of the CPAI.

The above finding is well described by Bashook, saying that assessor bias, or different interpretations of performance, increases when assessors

need to indicate how well a given task was performed (Bashook, 2005). That said, authors agree that, although assessor bias will never be eliminated, it can be greatly reduced by choosing appropriate assessors, as well as providing adequate training for the identified assessors (Bashook, 2005; PMETB Workplace based assessment Subcommittee, 2005; Tavares et al., 2013; Edens and Swoboda, 2013).

6.4 SUPPORT FOR THE IMPLEMENTATION OF THE UJ CPAI AS AN ASSESSMENT INSTRUMENT

The participants of the focus group discussion agreed that there remains a desire and need for the implementation of an instrument such as the UJ CPAI in the University Of Johannesburg Department Of Emergency Medical Care Bachelor of Health Sciences in Emergency Medical Care (UJ Dept. EMC BHS EMC) programme. the researcher will use this part of the chapter to discuss the support for implementation of the UJ CPAI in the UJ Dept. EMC BHS EMC programme.

As mentioned in Chapter two, credibility essentially refers to how well assessment results are supported by users thereof (Bashook, 2005). Also, for assessors to support the use of an assessment instrument, the instrument should be feasible. The researcher will now discuss the issues of credibility and feasibility of the UJ CPAI.

6.4.1 Credibility

In Chapter two, the researcher mentioned that educators and students who will use the assessments need to be supportive of the assessment method and value its use. Prober (2011) argues that the most important issue of credibility is the issue of transparency. This argument is supported by SAQA (Prober, 2011; SAQA, 2015). In order for assessment to be transparent, students and educators need to know exactly what criteria

students are being assessed on, and exactly how they will be awarded marks (Prober, 2011; SAQA, 2015). This was achieved in this research study by making the narrative section of the CPAI available to all of the students that would work in the UJ PRV during the data collection period, explained in Chapter three. Also, it was felt that the UJ CPAI was able to allow assessors to provide good feedback. Good feedback does not only improve transparency, but also identifies areas for remediation by the feedback provided to learners and module lecturers (PMETB WPBA sub-committee, 2005). This in turn, drives learning (Keating, Dalton, & Davidson, 2009).

In addition, there must be trust that assessment will not be biased and that the criteria will be applied fairly, objectively and consistently (Prober, 2011; SAQA, 2015). As shown in Chapter five, there were a number of participants who felt that that each narrative still needs to be interpreted. Also, in some cases, assessors will award marks according to their own set of beliefs and opinions regardless of what the narrative advises.

The argument is also made in Chapter two that assessment needs to be based on direct observation, rather than indirect reporting (Prober, 2011). This is important as it is well described that what students do in the classroom assessment opportunities such as simulations and OSCEs differ from what they really do in real-life situations (Rhethans et al., 2002; Tavares et al., 2013). The UJ CPAI was used to assess students' clinical performance while engaged with real-life work related duties and therefore the researcher argues that it was well-placed to assess clinical performance.

In summary, the researcher posits that making the narrative of the CPAI available to staff and students from the beginning and the perceived value of the feedback the CPAI allowed assessors to give improved the

transparency of the instrument. That said, the perceived bias associated with the allocation of marks could, in turn, detract from the credibility of the UJ CPAI. There must be trust that assessment will not be biased and that the criteria will be applied fairly, objectively and consistently (Prober, 2011; SAQA, 2015). If there is perceived bias associated with the awarding of marks, credibility will be negatively influenced.

6.4.2 Feasibility

In nearly all situations, feasibility becomes a question of available money, expertise, opportunity, resources, and time. (Joint Committee on Testing Practices, 1999; Linn, 1994). Users, such as clinical assessors and students, reflect greater satisfaction with tests that are short, easy to administer and less costly (Cadorin, Bagnasco, Rocco & Sasso, 2014). As could be seen from the responses in Chapter 5, there was overwhelming agreement that the UJ CPAI is too cumbersome and takes too long to complete in its current format. The researcher argues therefore that the UJ CPAI in its current format is not feasible for implementation. When the process of assessment becomes too long-winded and complicated, assessors may not complete the instruments as attentively as is necessary and may rush the process. This may result in the value of the assessment to be undermined (Joint Committee on Testing Practices, 1999; Linn, 1994).

6.5 RECOMMENDATIONS REGARDING THE FUTURE USE OF THE CPAI WITHIN THE EMERGENCY MEDICAL CARE SETTING.

The researcher showed that the validity, reliability, credibility and feasibility of the assessment instrument was low. The researcher identified a number of recommended changes to the UJ CPAI. These recommendations were related mostly to the format of the UJ CPAI, assessor training and the assessment process using the instrument. The

format of the assessment needed to be refined so that it provides clarity on the exact constructs that are assessed under each of the nine assessment areas. Also, the performance criteria for assessment areas one, two and three to be more focused on assessing the assessment area, as explained earlier in the chapter. The five-point scale used to award marks appears to be creating confusion and should be changed to a three-point scale.

Assessors should receive better training on the use of the UJ CPAI as an assessment instrument. This is may decrease assessor bias and improve the inter-rater reliability. Also recommended was the need to streamline the assessment process in order to make it easier to complete and less time-consuming.

Although these may seem to be simple changes, the process of changing the UJ CPAI is complex. Changing the UJ CPAI in a manner that is not informed by good evidence and that does not follow thorough scientific process, may cause the refined CPAI to have even lower validity and inter-rater reliability. The recommendations for further research in this area is discussed in Chapter 7.

6.6 SUMMARY

This Chapter presented and discussed some of the central findings that emerged from this study together with their potential implications for the refinement of the UJ CPAI. In summary, the researcher found that the construct validity of the UJ CPAI was low. Four areas were identified where the format of the UJ CPAI decreased construct validity. These four areas were: i) the narrative of assessment area one (preparation), ii) the narrative for assessment area two (professional presentation), iii) the narrative for assessment area three (history taking, communication and

general patient interaction) and, iv) the five-point scale used to award marks to the students.

The total number of Likert-type responses awarded on the questionnaire was 63 (seven participants awarding nine ratings each). Of the 63 ratings awarded, a good rating of four or five was awarded 78% of the time (n=49). This suggests that, overall, the participants agreed that the UJ CPAI had good content validity.

The inter-rater reliability was also found to be low, with the results from the 52 analysed cases showed that an α of more than 0.60 was found in only 67% of the cases it was used to assess. Three reasons are put forward that could have caused inter-rater reliability to be low. These were: i) low construct validity, ii) Inadequate assessor training on the use of the UJ CPAI and iii) assessor bias when awarding marks to students.

The researcher also argued that the UJ CPAI in its current format is not feasible for implementation. There was overwhelming agreement by the participants in the focus group discussion that the UJ CPAI is too cumbersome and takes too long to complete in its current format. The next and final Chapter summarises and concludes the study dealing with limitations, recommendations and suggestions for further study.

CHAPTER 7 – SUMMARY AND CONCLUSION

7.1 INTRODUCTION

This final Chapter provides a synopsis, conclusion and summary of this research study and dissertation.

7.2 SUMMARY OF THE RESEARCH STUDY AND DISSERTATION

In *Chapter 1*, the background and context to the study were introduced. The aim of the study was to critically investigate and describe the validity and reliability of the instrument for the assessment of clinical performance of emergency medical care students in the pre-hospital environment. A short introduction to the research methodology, problem statement, aims and objectives as well as the possible benefits was documented.

Chapter 2 introduced Emergency Medical Care education in South Africa, presented literature on assessment of clinical performance, and discussed an argument for the need to validate an instrument for the assessment of clinical performance in the South African pre-hospital emergency care education domain. This literature review was important as it attempted to frame the study within the context of existing knowledge.

In *Chapter 3* an argument for the use of a sequential mixed-methods design was made. The qualitative data emerged from a desktop study that analysed the format and structure of the UJ CPAI, responses from a self-designed, non-validated, evidence informed questionnaire and results from 104 completed marks sheets. This raw data was captured on a template. The findings of the quantitative data analysis are presented in *Chapter 4*. The emerging themes and trends identified from the focus group discussion are presented in *Chapter 5*. In *Chapter 6* the researcher

presented and discussed some of the central findings that emerged from this study together with their potential implications for the UJ CPAI. This final Chapter, *Chapter 7* provides a conclusion and overview of the research report.

In order to ensure validity, reliability and trustworthiness of this study, the researcher appropriately used a prospective, descriptive, sequential mixed methods design for this study. Three different research methods were used that followed a logical sequence. These were a literature review, quantitative data collection and focus group discussion following the method mentioned above. That was followed by thorough method triangulation of findings. Trustworthiness of the study was ensured by careful identification of discussion points for the focus group discussion, audio-recording of the focus group discussion and transcribing the discussion verbatim. Also, the researcher carefully triangulated and discussed the data using a well described method called method triangulation.

The main findings that emerged from this study were:

i) The construct validity for certain of the assessment areas was found to be poor. Possible reasons put forward for this were that the format of the assessment instrument did not enable assessors to adequately assess clinical performance. The results stemming from this study seemed to support the idea that the way the narrative was written meant that participants interpreted the narrative of the UJ CPAI differently. In addition, the researcher identified four assessment areas where the format of the UJ CPAI decreased construct validity. These four areas were: i) the narrative of assessment area one (preparation), ii) the narrative for assessment area two (professional presentation), iii) the narrative for assessment area three (history taking, communication and general patient

interaction) and, iv) the five-point scale used to award marks to the students.

ii) The content validity was found to be good for eight of the nine assessment areas. The area that showed poor content validity (assessment area three - history taking, communication and general patient interaction) also showed low construct validity. When discussing the construct validity of assessment area three, the focus group participants were of the opinion that the way the narrative for this assessment area was structured made it difficult to identify exactly what the student is being assessed on. It would therefore be expected that the participants awarded poor ratings for this assessment area in the questionnaire.

iii) The overall inter-rater reliability between assessors using the instrument was found to be low. That said, three areas were identified that caused the inter-rater reliability to be low. The most significant reason was the poor construct validity. The other two reasons were the need for improved assessor training, and finding ways to decrease assessor bias. The research indicated that different assessors interpret the same narrative very differently. This caused inter-rater reliability to be low.

iv) There was strong support noted from educators for the use of an instrument such as the UJ CPAI. The provisory was that the process needed to be streamlined and the instrument needed to be simplified since there was overwhelming agreement that the UJ CPAI is too cumbersome and takes too long to complete in its current format.

7.3 RECOMMENDATIONS

Considering the outcomes and findings of this study, the following recommendations are made:

- The format of the assessment needs to be refined so that it provides clarity on the exact constructs that are assessed under each of the nine assessment outcomes.
- The assessment criterion for assessment outcomes one, two and three to be more focused on assessing the assessment outcome, as explained in Chapter 6.
- The five-point rating scale should also be changed to a three-point rating scale.
- There should be improved training given to the assessors on the use of the UJ CPAI as an assessment instrument. This is thought to decrease assessor bias and improve the inter-rater reliability of the results.
- Also recommended was the need to streamline the assessment process in order to make it easier to complete and less time-consuming. This meant making changes such as simplifying the rating scale to a three point rating scale, as well as simplifying the narrative for the assessment areas.

Although these may seem to be simple changes, the process of changing the UJ CPAI remains complex. Changing the UJ CPAI in a manner that is not informed by good evidence and that does not follow a scientific process, may cause the refined CPAI to have even lower validity and inter-rater reliability.

7.4 POTENTIAL CONTRIBUTION OF THE STUDY

This study produced new literature on the assessment of South African Emergency Medical Care students in the prehospital environment. This study also made use of an alternative method to investigate the content as well as construct validity of newly developed assessment instruments. The study highlighted that the current UJ CPAI has areas that required refinement and suggests evidence informed changes to the instrument are necessary in order to improve its validity and reliability as an assessment instrument. This study highlighted the need for an instrument such as the UJ CPAI and may lead to the researcher embarking on further study to further improve and refine a model for the implement the UJ CPAI.

7.5 LIMITATIONS

With all studies, there are limitations and obstacles that are encountered which may impact on the validity and reliability of findings. In this study, the researcher wishes to acknowledge the following challenges and limitations:

The limited published literature on validated instruments to measure clinical performance in the pre-hospital setting in general but more specifically in the local South Africa context was a challenge faced in completing this study.

The number of emergency calls received by the UJ departmental vehicle was lower than expected over the study period. Both the researcher and supervisor worked on the Departmental vehicle; consequently CPAI forms that they completed could not be included in the sample of forms that were included in study. Having said this, the 52 data set recorded were seen as sufficient in number to determine inter-rater reliability. Factors

that played a role were the vehicle being involved in an accident towards the end of the year, as well as the crew of the vehicle being the victims of an armed robbery whilst on shift.

The number of participant that were included in this study was small. That was because the inclusion criterion included only academic staff, clinical facilitators and external assessors that had used the CPAI to assess students on the UJ Primary Response Vehicle (PRV). The reason for this inclusion criterion was that these were the only individuals that would have used the CPAI in the setting it was intended to be used and therefore had sufficient exposure, experience and insight into the instrument to make meaningful comment regarding it's content validity. The limited sample size and the narrow focus of this study did not lend the results to be generalizable to pre-hospital emergency care education as a whole.

7.6 PROPOSALS FOR FURTHER STUDIES

The researcher is planning to proceed with a PhD study, making required changes, refining the CPAI and piloting the refined UJ CPAI in a multi-centered study involving Emergency Medical Care programmes at the University of Johannesburg Department EMC and other HEIs engaging in the education and training of ECPs.

This study may also lead to the development and implementation of an instrument used to assess simulation-based assessments within the Professional Degree as well as Diploma in Emergency Medical Care Programmes.

This research showed the need for adequate assessor training prior to the use of an assessment instrument. This may inform the development of a

formalised assessor training event that is registered through the continuing professional development (CPD) office.

7.7 CONCLUSION

In this study, the researcher made use of a number of different methods in order to answer the research question. Since the investigation into the validity and reliability of this assessment instrument was performed in 'uncontrolled' environments, common statistical tests such as the calculation of interclass correlation coefficients across a larger sample of assessors and students were not possible. The researcher therefore needed to employ other less commonly described tests, such as the calculation of correlation coefficient for two raters assessing a single student for a large number of assessments. Also used in this study was a mixed-methods methodology. The researcher employed a variety of research methods. These included a literature review, administration of a survey questionnaire assessing the construct and content validity of the CPAI, analysis of quantitative data gathered from completed UJ CPAI mark sheets, analysis of qualitative data generated by a focus group discussion and method triangulation and corroboration of quantitative and qualitative data sets. This methodology, although well described, has not yet been widely used for the investigation of validity and reliability instruments for the assessment of clinical performance of emergency medical care students in the pre-hospital environment.

This study highlighted some interesting similarities and differences in the assessment instruments used for the assessment of clinical performance. Central findings included that the construct validity of the CPAI needs to be improved, that the content validity is shown to be satisfactory and that the instrument was capable of producing results showing good inter-rater reliability more than half the time. The study shows support from

educators for the implementation of an instrument such as the UJ CPAI; however, in its current format the assessment process appears too cumbersome. Refinements and further study is needed to develop the instrument to a point where it can be considered scientifically validated and universally accepted as suitable for the assessment of clinical competence in the prehospital emergency care environments.

REFERENCES

Andrew, S., Halcomb, E.J. 2009. *Mixed Methods Research for Nursing and the Health Sciences*. Oxford: Wiley – Blackwell.

Ausubel D.P., Novak J.D. & Hanesian H. (1978) *Educational Psychology: A Cognitive View*. Holt, Rinehart & Winston, New York.

Bashook, P.G. 2005. Best Practices for Assessing Competence and Performance of the Behavioral Health Workforce. *Administration and Policy in Mental Health*. 32 (5/6): 563-589

Beckman, J., Cook, D., & Mandrekar, J. (2005). What is the Validity Evidence of Assessments in Clinical Teaching? *Journal of General Internal Medicine*, 20(12):1159-1164

Bickley, L.S., Szilagy, P.G. 2003. *Bates' Guide to Physical Examination and History Taking*, 8th ed. Philadelphia: Lippincott, Williams and Wilkins.

Boud, D., Falchikov, M. 2007. *Rethinking Assessment in Higher Education: Learning for the Longer Term*. New York: Routledge.

Bouriscot, K., Etheridge, L., Ker, J., Sambadam, E., Setna, Z., Smee, S., Sturrock, A. Preliminary Report with draft Consensus Statements and Recommendations for the Performance Assessment Theme. Ottawa Conference 2010. Ottawa

Brink, H., van der Walt, C., van Rensburg, G. 2012, *Fundamentals of Research Methodology for Healthcare Professionals*. 3rd ed. Cape Town: Juta

Carraciou, C.L., Benson, B.J., Nixon, J., Derstine, P.L. 2008. *From the Educational Bench to the Clinical Bedside: Translating the Dreyfus*

Developmental Model to the Learning of Clinical Skills. *Academic Medicine*. 83 (8):761-767

Christopher, L.D. 2007. An investigation into the non-compliance of Advanced Life Support Practitioners with the guidelines and protocols of the Professional Board for Emergency Care Practitioners. Masters Dissertation. Durban: Durban University of Technology.

Cook, D., Beckman, T. 2006. Current Concepts in Validity and reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine*, 119: 166e7– 166e.16

Copeland, H. & Hewson, M., (2000). Developing and Testing an Instrument to Measure the Effectiveness of Clinical Teaching in an Academic Medical Center. *Academic Medicine*, 75(2):161-166

Council on Higher Education. 2011. Work Integrated Learning: A Good Practice Guide. Pretoria: IeCommunications.

Cresswell J.W. (2009). Research Design Qualitative, Quantitative and Mixed-Methods Approaches. California: Sage Publications

De Sanctis, F. 2006. Theories and Techniques of Learning. Napoli: Tecnodid Editrice.

Deakin University. 1994. A Correlation of Readings Related to Competency Based Training. Deakin University, Australia.

Development of Clinical Reasoning in Simulation. *Journal of Nursing Educator*, 51(6):326-333.

Dreifuerst, K.T. 2012. Using Debriefing for Meaningful Learning to Foster to foster Development of Clinical Reasoning in Simulation. *Journal of Nursing Educator*, 51(6):326-333.

Dreyfuss, S.E. 2004. The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science Technology & Society* 2004 (24):177

Dreyfuss, S.E. 2004. The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science Technology & Society* 2004 (24):177

Dreyfuss, S.E., Dreyfuss, H.L. 1980. *A Five-Stage Model Of The Mental Activities Involved In Directed Skill Acquisition*. Research Report Submitted to the ASFC, USAF Under Contract to the University of California. USA: United States Government.

Dreyfuss, S.E., Dreyfuss, H.L. 1980. *A Five-Stage Model Of The Mental Activities Involved In Directed Skill Acquisition*. Research Report Submitted to the ASFC, USAF Under Contract to the University of California. USA: United States Government.

Edens, M., Swoboda, T.K. 2013. Development of Workplace-Based Assessment Tools for Graduate Training Programmes: a 360 Degree Evaluation Model. Presented at the International Conference on Residency Education. Royal College of Physicians and Surgeons of Canada. 17 September 2013.

Fink LD (2003). *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. Jossey-Bass, San Francisco.

Fink LD (2007). The Power of Course Design to Increase Student Engagement and Learning. Winter: 13-17.

Fink M (2005). *Conducting Research Literature Reviews: From the Internet to Paper*, 2nd ed. Sage, London.

Flemke K. 2004. *Methodology: Overview of the Research Design*, Ch 3 *Incarcerated women's experience of rage: A phenomenological understanding*. <http://scholar.lib.vt.edu/theses/available/etd-03132003-224950/unrestricted/04KFlemke-Chpt3methods.pdf>

Freeman, R., Lewis, R. 1998. *Planning and Implementing Assessment*. London: Kogan Page

Gaberson, K.B., Oemermann, M.H., Shellenbarger, T. 2015. *Clinical Teaching Strategies in Nursing*. Fourth ed. USA NY: Springer

Hall-Ellis, S.D., Grealy, D.S. 2013. The Dreyfus Model of Skill Acquisition: A Career development Framework for Succession Planning and Management in Academic Libraries. *College and Research Libraries*, 2013:587-601.

Heinrich, B., Heinrich, M. 1984. The pit-trapping foraging strategy of the ant lion, *Myrmeleon immaculatus* DeGeer (Neuroptera: Myrmeleontidae). *Behavioural Ecology and Sociobiology* 1984 (14):151-160

Hill, A.P., Freeman, G.K. 2011. *Promoting Continuity of Care in General Practice: An RCGP Policy Paper*. London: Royal College of General Practitioners.

HPCSA (Health Professions Council of South Africa), PBEC (Professional Board for Emergency Care) 2011. *Capabilities of Emergency Care Providers*: June 2011. Available from: http://www.hpcsa.co.za/Uploads/editor/UserFiles/downloads/emergency_c

[are/capabilities_baa_eae_eep_ect_jul_2013_final.pdf](#). Accessed 20 June, 2015.

HPCSA (Health Professions Council of South Africa). 2005. EC News: Newsletter of the Professional Board for Emergency Care. Pretoria: HPCSA

HPCSA (Health Professions Council of South Africa). 2006. EC News: Newsletter of the Professional Board for Emergency Care. Pretoria: HPCSA

HPCSA (Health Professions Council of South Africa). 2010. Professional Board for Emergency Care. <http://www.hpcsa.co.za>. Accessed on 15 June 2015.

Independent Police investigative Directorate. Batho Pele Principles. Available from: http://www.icd.gov.za/about%20us/batho_pele.asp. Accessed 29 June 2015.

Keating, J., Dalton, M., & Davidson, M. (2009). Assessment in clinical education. In C. Delany, & E. Molloy (Eds.), *Clinical education in the health professions* (First ed.). Chatswood: Elsevier Australia.

Kieffer, V., Davies, K., Gibson, C, Middleton, M., Munday, J., Shalets, S., *et.al.* 2015. Society for Endocrinology Competency Framework for Adult Endocrine Nursing: 2nd edition. *Endocrine Connections*. 2015 (4):1-17.

Kogan, J.R., Hauer, K.E. 2006. Brief Report: Use of the Mini Clinical Evaluation Exercise in Internal Medicine Core Clerkships. *Journal of General Internal Medicine*. 21(5):501-502.

Kovacs, G., Law, J.A. 2008. *Airway Management in Emergencies*. USA: McGraw-Hill.

Krueger, R. 2015. Analyzing Focus Group Results. In Krueger, R., Casey, M. Focus Groups: A Practical guide for applied Research. California: Sage Publications.

Landis, J., Koch,G. (1977). Measurement of observer agreement of categorical data. *Biometrics*, 33:159-174

Lewis, R. 2011. Interprofessional learning in acute care: Developing a theoretical framework. *Nurse Education Today*. 32:241-245

Lyon, L.J. 2015. Development of teaching expertise viewed through the Dreyfus model of skill acquisition. *Journal of the Scholarship of Teaching and Learning*. 15 (1): 88 - 105.

Mamelok, J. 2009. Workplace-Based assessment portfolios in Licensing for General Practice Speciality Training. *Education for Primary Care*. 20:139-142

Martin, A., Rees, M., Edwards, M. 2011. *Work-Integrated Learning: A Template for Good Practice*. Wellington: Ako Aotearora.

Maxwell, J. 2005. Qualitative Research Design: An Interactive Approach. 3rd Ed. Los Angeles: Sage Publications

Melville, A., Johnston, C. 1982. *Cured to Death*. Secker and Warburg, London.

Mitchell, C., Bhat, S., Herbert, A., Baker, P. 2011. Workplace Based Assessments of Junior Doctors: Do Scores Predict Training Difficulties? *Medical Education* 45:1190-1198

Morgan, L.D. 2007. Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, 1(1):48-76

Norusis, M. 2003. *SPSS 12.0: Statistical Procedures Companion*. New Jersey: Prentice Hall.

Patton, M. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34(5): 1189–1208

PMETB Workplace based assessment Subcommittee. (2005). Workplace Based Assessment. A paper from the PMETB Workplace Based Assessment Subcommittee January 2005.

Positivistic Inquiry Paradigm versus Naturalistic Inquiry Paradigm. *Journal of Medical Education*, 5(2):75-80

Republic of South Africa. 2013. Department of Higher Education and Training. White Paper for Post-School Education and Training — building an expanded, effective and integrated post-school system. Pretoria.

Rethans J., Norcini JJ., Baron-Maldonado M., Blackmore D., Jolly BC., LaDuca T., Lew S., Page GG., & Southgate LH. 2002. The relationship between competence and performance: Implications for assessing practice performance. *Medical Education*, 36, 901-909

Riegelman, R. 2005. *Studying a study and Testing a Test: Reading Evidence-Based health research*. Philadelphia: Lippincott, Williams and Wilkins.

Roberts, T. (2011) Assessing professionalism: The biggest challenge of all. Chapter from Assessment and training of tomorrow's Doctors: 3rd Beijing international forum of medical deans. PMPH, USA

Rosental, R., Westen, D. 2003. Quantifying Construct Validity: Two Simple Measures. *Journal of Personality and Social Psychology*. 2003 (84): 608–618.

Salkind, N. 2004. *Statistics for People Who (think they) hate statistics*. Kansas: Sage.

Sanders, M.J., McKenna, K., Quick, G., Lewis, L.M. 2001. *Mosby's Paramedic Textbook*. Missouri: Mosby.

Scotland, J. 2012. Exploring the Philosophical Underpinnings of Research: Relating Ontology and Epistemology to the Methodology and Methods of the Scientific, Interpretive, and Critical Research Paradigms. *English Language Teaching* 5(9):9-16

Sireci, S., Faulkner – Bond, M. 2014. Validity evidence based on test content. *Psicothema*,26(1). 100-107

Skinner, D. 2007. Qualitative Methodology: An Introduction. In Joubert, G & Ehrlich, R. (Eds.) *Epidemiology: A research manual for South Africa*. Cape Town: Oxford publications.

Stanley MJC, Dougherty JP (2010). A Paradigm Shift in Nursing Education: A New Model. *Nursing Education Perspectives* 31(6): 378-380.

Stathers, C.L., Pecora, D., Copeland, D.L., Woodforth, B. 2011. Introduction to Critical Care Transport. In Murphy, M., Stathers, C.L.,

Pecora, D., McEvoy, M., Jeffery, S.R. 2011. *Critical Care Transport*. Studbury, MA: Jones and Bartlett.

Stein, C. 2010. Pre-Hospital Emergency Care Student Experience with Paediatric Cases in Johannesburg, Gauteng. Masters Dissertation. Johannesburg: University of the Witwatersrand.

Takavol, M., Zeinaloo, A. (2004). Medical Research Paradigms: Positivistic Inquiry Paradigm versus Naturalistic Inquiry Paradigm. *Journal of medical education* 5(2):75-80

Tavares, W., Boet, S., Theriault, R., Mallette, T., Eva, K.V. 2013a. Global Rating Scale for the Assessment of Paramedic Clinical Competence. *Prehospital Emergency Care*. January / March:57-67

Tavares, W., LeBlanc, V.R., Mausz, J., Sun, V., Eva, K.W. 2013b. Simulation-Based Assessments of Paramedics and Performance in Real Clinical Contexts. *Prehospital Emergency Care*. July / September:1-7

Tavares, W., LeBlanc, V.R., Mausz, J., Sun, V., Eva, K.W. 2013b. Simulation-Based Assessments of Paramedics and Performance in Real Clinical Contexts. *Prehospital Emergency Care*. July / September:1-7

United Kingdom Postgraduate Medical Education Training Board Workplace based assessment Subcommittee. (2005). Workplace Based Assessment. A paper from the PMETB Workplace Based Assessment Subcommittee January 2005.

University of Johannesburg Academic Development and Support. Work Integrated Learning and Service Learning policy. University of

Johannesburg; 2013. Available at:

<http://www.che.ac.za/sites/default/files/publications/White%20Paper%20-%20final%20for%20web.pdf>

Van Tonder, B. Learning Guide for Emergency Medical Care Clinical practice III. Second edition. University of Johannesburg; 2015.

Van Tonder, B. Learning Guide for Emergency Medical Care Clinical practice III. Second edition. University of Johannesburg; 2015.

Vincent-Lambert, C. (2011). A Framework for Articulation between the Emergency Care Technician Certificate and the Emergency Medical Care Professional Degree; Doctoral Theses, University of Free State, <http://etd.uovs.ac.za/ETD-db/theses/available/etd-08232012-152423/unrestricted/Vincent-LambertC.pdf> accessed April 2015.

Vincent-Lambert, C. 2015. International Perspectives: South African Ambulance Services 2020. In: Wankhade, P., Makway-Jones, K. *ed. Ambulance Services: Leadership and management perspectives*. London: Springer.

Wass, V., Van Der Vleuten, C., Shatzer, J., Jones, R. (2001). Assessment of clinical competence. *Lancet*, 357(9260): 945-949.

Yorke, M. 2005. Issues in the Assessment of Practice-Based Professional Learning. A Report Prepared for the Practice-Based Professional Learning CETL at the Open University

ANNEXURE 1 – INVITATION TO PARTICIPATE IN THE STUDY

19 February 2015

Invitation to Participate in a Masters Research Project

Good day,

My name is Bernard van Tonder. I am currently registered with the Cape Peninsula University of Technology for a Master of Emergency Medical Care, which is a full research Masters. For the qualification, I am conducting research into the validity of the newly developed workplace based clinical performance assessment instrument at the University of Johannesburg. My research topic is **AN INVESTIGATION INTO THE VALIDITY AND RELIABILITY OF AN INSTRUMENT FOR THE ASSESSMENT OF CLINICAL PERFORMANCE DURING WORK INTEGRATED LEARNING OF EMERGENCY MEDICAL CARE STUDENTS AT THE UNIVERSITY OF JOHANNESBURG.**

I would like to invite you to participate in the study. I would like to invite you to take part in research step four, where the newly developed UJ Clinical Practice Assessment Instrument (UJ CPAI) will be used on the Departmental Primary Response Vehicles. For this phase of the study, you will be assessed by academic staff and external assessors using the UJ CPAI while you are performing clinical duties. Please refer to the information sheet "how to complete the UJ CPAI" for more information.

Participation in the study is voluntary and the decision not to participate will be respected. You will be allowed to withdraw at any point if you do not want to continue with the study without incurring any penalties or any form of retribution. There will be no direct benefit, monetary, extra mark

allocation or otherwise, from participating in the study. Participation in the study will also not require from you to incur monetary loss.

All of the data gathered will be treated as confidential. Only the supervisors and I will have access to the raw data. The personal information and results will not be distributed in any way.

Please feel free to ask any questions relating to the study to the person explaining it to you. Should you wish to participate, please indicate such below and return to the researcher.

Yours truly,

Bernard van Tonder

Student: Master of Emergency Medical Care

Cape Peninsula University of Technology

Department of Emergency Medical Sciences

212306189

079 874 2368

.....
Consent

I,..... have read the invitation to take part in the research study. I understand that taking part in the study will not result in direct benefit such as extra mark allocation or monetary gain. I also understand that should I choose not to take part in the study, I will suffer no retribution in any way. I hereby consent / do not consent to take part in the study.

SIGNATURE

DATE

ANNEXURE 2 – INFORMATION SHEET ON HOW TO COMPLETE THE UJ CPAI



DEPARTMENT OF EMERGENCY MEDICAL CARE CLINICAL PERFORMANCE ASSESSMENT GUIDELINES FOR COMPLETION OF THE CLINICAL PRACTICE EVALUATION INSTRUMENT

Step 1 - The clinical performance assessment instrument narrative is available to the staff and students in the black file on R47.

Step 2 - While the student that is being assessed perform the required duties and procedures in order to manage the incident, the clinical assessors will take cognisance of how the student performs and provide guidance where required. If the scenario permits, the assessor may already start to complete the clinical assessment feedback form. The needs of the patient may not be neglected as the focus of the assessment is based on how well students are able to care for patients.

Step 3 - After the patient has been handed over to the relevant personnel, such as ambulance personnel or emergency department staff, the student will complete a University of Johannesburg Patient Care Record (PCR) as well as the student self-assessment component of the clinical assessment feedback form.

Step 4 – The incident will be debriefed and the student PCR scrutinised. Each assessor will complete sections 7 and 9 the assessor feedback section of **their own** feedback form, and may not discuss their feedback with each other. The marks awarded to the student during the incident may not be altered during the debrief.

Step 5 – After the documentation has been completed, the forms will be placed in the red file. The forms of each incident will be filed separate from forms for other incidents. The file dividers can be used for this.

If there is anything that is unclear or requires clarification, or if you need any more information, please contact me on 011 559 6248 or 079 874 2368.

Bernard van Tonder

Student: Master of Emergency Medical Care

Cape Peninsula University of Technology

Department of Emergency Medical Sciences

212306189

079 874 2368

ANNEXURE 3 – UJ CPAI NARRATIVE



DEPARTMENT OF EMERGENCY MEDICAL CARE

GUIDELINES AND CRITERIA FOR STUDENT CLINICAL ASSESSMENT

This document provides a guide to the consistent grading of students performance in the nine critical areas of clinical interaction. Please note this document should be consulted prior to and during the awarding of points for each category.

Date of review: January 2015

Assessment Area 1: Preparation

Rating and Criteria

1 = Does not comply

Student arrives late. The necessary equipment is not present and / or functional. Vehicle and associated equipment is dirty. There is evidence of used medical and surgical packaging / equipment from previous calls. The vehicle check sheet has not been properly completed. The equipment check sheet has not been properly completed.

2 = Needs Improvement

The vehicle and equipment check sheets have been properly completed. Student is on time, most of the equipment required is in place, the vehicle and equipment is clean. However, the required levels of stock are not present. The rear of the vehicle not properly packed. Student does not appear to know exactly where items of equipment are when they are needed. The shift is delayed due to the student having to fetch a few items that should have been organised beforehand from the store.

3 = Meets Minimum Standards

Student is on time. The vehicle and equipment check sheets have been properly completed. All of the equipment required is in place, the vehicle and equipment is clean and functional. The equipment is well positioned and easily located.

4 = Exceeds Minimum Standards

Student reports ahead of time. The vehicle and equipment check sheets have been properly completed. The vehicle and equipment are immaculate. The vehicles tyres are polished. The windscreen is 100% free of smudge marks and residues. The equipment is well positioned, quickly located and easily accessible.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

Assessment Area 2: Professional Presentation

1= Does not comply

Student is not properly dressed for professional clinical interaction e.g. clothes are dirty and / or creased, uniform is incomplete no name badge or rank insignia. The student's hair is dirty / greasy. The student's nails are too long and or are dirty. The student's hair is unkempt. Their shoes are inappropriate and or are dirty. Items of PPE are missing or damaged.

2 = Needs Improvement

Student's personal appearance is acceptable however their clothing and PPE does not comply with the minimum standards of the department as per the departmental rules and regulations.

3 = Meets Minimum Standards

Student presents the image of a health care professional. Their clothing is clean and smart and compliant with the uniform and PPE requirements of the department and profession. Their personal hygiene and appearance is that expected of a health care professional.

4 = Exceeds Minimum Standards

Student is extremely well dressed and groomed. Their shoes are polished and or are very clean. Their uniform is immaculate. Student presents the image of a health care professional.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

**Assessment Area 3: History Taking, Communication and General
Patient interaction**

1 = Does not comply

Student appears anxious and shows a lack of confidence when trying to interact with the patient. Their non-verbal body language is poor. They do not position themselves properly when engaging with the patient. Eye contact is avoided. The student appears too shy and does not come across as a confident health care professional. The student speaks too fast or too loudly or too softly and often repeats questions. The history taken lacks structure and direction. Many inappropriate questions are asked. The history taking is repetitive, shallow and does not relate clearly to the patient's chief complaint. Major important areas of the medical history are missed e.g. allergies. The patient is not made to feel at ease and is not properly included in the consultation process. The supervising clinician has to frequently interrupt and or clarify issues for the student and patient.

2 = Needs Improvement

The student is nervous and appears anxious; however they are able to interact with the patient. The supervising clinician has to (from time to time) remind the student of areas that they have missed. Their bedside manner and non-verbal body language needs work. They do however position themselves properly when engaging with the patient, eye contact is acceptable. The history taken has some structure and direction, however a few inappropriate questions are asked. Most of the important areas of patient history are covered e.g. allergies etc.

3 = Meets Minimum Standards

The student appears relaxed and is able to confidently interact with the patient. The supervising clinician does not need to step in and remind the student of issues that they have missed. Their bedside manner and non-verbal body language is acceptable. They position themselves properly when engaging with the patient, eye contact is acceptable. The student speaks with a reasonable pace and volume. The history taken has structure and direction with all the appropriate questions asked. The important areas of patient history are covered e.g. allergies etc. The patient understands the questions as they are clear and concise.

4 = Exceeds Minimum Standards

The student is calm and confident and comes across as a knowledgeable health care professional. Their bedside manner and non-verbal body language is good and the patient appears relaxed and at ease and seems to trust the student's judgement. The history taken is very well structured and directed to quickly and efficiently obtain all the appropriate information. The supervising clinician does not have to intervene in any way with the history taking.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

Assessment Area 4: Clinical Assessment and Diagnostics Skills

1= Does not comply

The student appears uncertain about which assessments and / or diagnostic tests are to be conducted. Important assessments are omitted. The student is unable to make a diagnosis or makes an incorrect diagnosis which they are unable to substantiate. The student is unable to link their theoretical knowledge (Anatomy, Physiology and General Pathology) to their clinical assessments and associated findings. When examining the patient the student appears anxious and uncertain and does not complete the necessary assessments properly or performs unnecessary assessments that are not indicated for the patient's condition. Clinical assessments that are conducted are incorrectly performed. The student takes too long to complete the clinical assessment.

2 = Needs improvement

The student knows which assessments and or diagnostics tests are to be conducted. The student is able to make a suitable diagnosis but cannot fully substantiate the diagnosis by linking their theoretical knowledge (Anatomy, Physiology and General Pathology) to their clinical assessments and subsequent findings. When examining the patient the student appears anxious and uncertain and does not complete the necessary assessments properly. The student is slow.

3=Meets Minimum Standards

The student knows which assessments and or diagnostics tests are to be conducted. The student is able to make a suitable diagnosis and can fully substantiate their diagnosis by linking their theoretical

knowledge (Anatomy, Physiology and General Pathology) to their clinical assessments and subsequent findings. When examining the patient the student completes the necessary assessments properly. The assessments are carried out within a reasonable time frame.

4=Exceeds Minimum Standards

The student knows which assessments and or diagnostics tests are to be conducted. The student is able to make a suitable diagnosis and can fully substantiate their diagnosis by linking their theoretical knowledge (Anatomy, Physiology and General Pathology) to their clinical assessments and subsequent findings. When examining the patient the student appears calm and confident, all of the necessary assessments are conducted properly. The student conducts the physical assessments quickly, smoothly and efficiently. Their bedside manner and non-verbal body language is good and the patient feels relaxed and at ease and obviously trusts the students judgement.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

Assessment Area 5: Decision-Making relating to clinical interventions

1= Does not comply

The student is unable to correctly formulate a treatment plan. There is no evidence-based practice evident in the selected treatment plan. The student is not able to prescribe the correct medication/s, and the overall knowledge of pharmacology is poor. Overall the decisions made by the student are inappropriate and do not address the patient's chief complaint adequately, nor were all the patient factors taken into consideration. The supervising clinician has to regularly intervene and / or take over to ensure that a suitable treatment plan is formulated. The treatment plan is not discussed with the patient. The student is not able to explain to the patient what their diagnosis is nor can they explain and / or validate the proposed treatment plan. Incorrect use of medical terminology is noted.

2 = Needs improvement

The student is able to formulate a treatment plan with the assistance of the supervising clinician. The student cannot fully justify decisions about planned interventions against evidence-based practice. The student can prescribe the correct medication however the overall knowledge of pharmacology requires some revision. The student decided on necessary interventions but is unable to justify the rationale behind the interventions. The student is able to formulate a basic management plan however certain components are missing. The supervising clinician has to help the student to moderate and or adjust the initial treatment plan. The treatment plan is discussed with the patient but the patient is not fully satisfied nor does the patient understand what is to be done. The student struggles to explain to the

patient their diagnosis or what treatment plan entails. Student requires assistance from clinician.

3=Meets Minimum Standards

The student is able to make appropriate clinical decisions based on evidence based best practice and can formulate and validate their treatment plan. Clinical procedures and or interventions are recommended taking the patients diagnosis, presentation and context in relation to distance and time from receiving facility into consideration. The correct medications are prescribed. The student is able to explain to the patient in simple understandable “layman’s” language what their diagnosis / condition is and what their treatment plan entails. The student is able to properly answer the patient’s questions.

4=Exceeds Minimum Standards

The student makes sound clinical decisions that are reflective of evidence based best practices that are linked to the history taken and finding from the clinical assessment conducted. There is a holistic approach to decisions about treatment options. The student prescribes the correct medication; pharmacological knowledge is good and student is able to also suggest alternatives. The treatment plan is discussed with the patient and adapted to suit the patient’s needs. The patient is happy and satisfied. The student is able to articulately explain to the patient in “layman’s” terms what the condition is and what their treatment plan entails. The student can answer the patient’s questions succinctly with a good theoretical knowledge.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

<p style="text-align: center;">Assessment Area 6: Clinical, Procedures, Skills and Treatment Interventions</p>

1= Does not comply

The student's manual dexterity skills are poor. Incorrect preparation and or use of equipment are noted. Poor attention to infection control is noted. Student practices unsafely jeopardising the patient, themselves and or observers. The incorrect equipment is selected or utilised to complete a procedure. The treatment modalities are used incorrectly or not appropriately for the condition. Wound assessment, dressing and general management of wounds are poor. The student lacks finesse and is uncertain how to perform clinical procedures. The student takes very long to complete procedures. Student requires assistance from supervising clinician to complete procedures. Consent has not been obtained from patient for procedures performed. The student does not involve the patient during the treatment. The student handles the patient in a rough manner. The patient is not given any advice by the student. The patient is not managed in an empathetic, caring and / or professional manner.

2 = Needs improvement

The student is able to complete the treatments / procedures however manual skills need to be improved. Correct preparation and or use of equipment are evident. The student requires assistance from the supervising clinician. Infection control is considered, however the student needs to improve to ensure that they practice safely. Health and safety of both patient and student is taken into consideration. The student works moderately paced however still needs to improve on time management. Consent is obtained from patient. The student attempts to involve the patient during the treatment however has to improve patient participation. Student makes an attempt to give patient advice.

3=Meets Minimum Standards

The student is able to complete the treatment / procedures within an acceptable time using the correct technique and equipment. Infection control principles are adhered to. The student takes Health and Safety of patients, others and themselves into consideration. Wound management is effectively carried out following best practice guidelines. Consent is obtained from patient and patient is totally involved in treatment. The student gives the patient appropriate advice and emotional support.

4=Exceeds Minimum Standards

The student is able to smoothly, quickly and effectively complete the treatment / procedures within an acceptable time using the correct equipment and technique. Student's dexterity and manual skills are well honed and to the eye of the observer denotes expertise. Infection control principles are strictly adhered to. The student in their

performance of clinical procedures is noted to be excellent. The student obtains consent to treatment; the patient is well briefed on the treatment details at every step of the way. The student takes time to calm and reassure the patient whilst carefully explains the clinical procedures to the patient.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

Assessment Area 7: Clinical Documentation and Record keeping

1= Does not comply

The student's documentation and PRF writing is haphazard and does not follow a logical sequence for e.g. Chief complaint, History, on examination, etc. Vital pieces of information relating to history, vital signs, times, medications, assessments and treatments are omitted. A black ballpoint pen is not always used. The writing is untidy and illegible. There is gross spelling and grammatical errors of commonly used words and medical terms. There is incorrect use of Medical terminology. Supervising clinician often has to help the student to re-write clinical notes. The overall impression left is that the student did not care when writing up the documentation. The report is not written up immediately after the patient is seen but some hours after

treatment. Patient consent to treatment or refusal of transport is not documented. Advice given to patient is not documented.

2 = Needs improvement

The student is able to write a medical report / PRF following a logical sequence however there are still pieces of information that are missing which the supervising clinicians has to remind the student about. A black ballpoint pen is used for record-keeping. The handwriting can be improved upon. There are a few spelling and grammatical errors and Medical terminology is used incorrectly. The supervising clinician sometimes has to help the student make alterations and corrections to the record-keeping. Patient consent to treatment is documented.

3=Meets Minimum Standards

The student is able to accurately record all relevant information following a logical sequence. File / PRF is well written and neatly presented using a black ballpoint pen. Correct use of Medical terminology, few spelling and grammatical errors are noted. Patient consent to treatment is documented together with advice given to patient.

4=Exceeds Minimum Standards

The student meticulously records all information relating to the consultation and treatment of the patient using a black ballpoint pen. Patient consent to treatment documented. No spelling and grammatical errors, correct use of Medical terminology. Student is able to write a succinct medical report. All patient questions and or concerns relating to treatment and outcome are clearly documented. Student accurately

records the time and place of treatment. Overall presentation of file / PRF is neat and legible and professional looking.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

Assessment Area 8: Handover

1 = Does not comply

Student appears very anxious and shows a lack of confidence when trying to interact with the doctor and / or emergency department staff. Their non-verbal body language is poor. They do not position themselves properly when engaging with the doctor. Eye contact is avoided. The student appears shy and does not come across as a confident health care professional. The student speaks too fast, too loudly or too softly and often has to repeat information. The handover lacks structure and direction. Many inappropriate points are mentioned. The handover is repetitive and does not relate clearly to the patient's chief complaint and subsequent management. Major important areas of patient history and or management are missed e.g. allergies etc. The supervising clinician has to frequently interrupt and or clarify issues for the student.

2 = Needs Improvement

The student is nervous and appears anxious; however they are able to interact with the doctor and or emergency department staff. The supervising clinician has to from time to time remind the student of issues that they have missed. Their bedside manner and non-verbal body language needs work. They do position themselves properly when engaging with the doctor, eye contact is acceptable. The student speaks too fast or too loudly or too softly and often repeats information. The handover has some structure and direction, however some inappropriate information is provided. Most of the important areas of patient history and subsequent management are covered e.g. allergies etc.

3 = Meets Minimum Standards

The student appears relaxed and is able to confidently interact with the doctor and or emergency department staff. The supervising clinician does not need to step in and remind the student of issues that they have missed. Their bedside manner and non-verbal body language is acceptable. They position themselves properly when engaging with the doctor, eye contact is acceptable. The student speaks at a reasonable pace. The handover provided has structure and direction with all the appropriate information being provided. The important areas of patient history and subsequent management are covered e.g. allergies etc.

4 = Exceeds Minimum Standards

The student is calm and confident and comes across as a knowledgeable and professional member of the health care team. Their bedside manner and non-verbal body language is good and the doctor is attentive and seems to trust the student's judgement. The

handover is logical, succinct and very well structured to quickly and efficiently provide all the appropriate information. The supervising clinician does not have to intervene in any way with the handover.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

<p>Assessment Area 9: Ability to reflect on own practice, self-critique and accept clinical direction</p>
--

1= Does not comply

The student lacks the ability to critically analyse their interaction and treatment of the patient. There is no evidence of self-critique. The student is unable to accept critique and advice from the supervising clinician. The student does not accept clinical direction and will often argue with the clinician whilst consulting with the patient. The student may also ignore the patient's requests or critique. The student is often defensive and may sometimes come across as aggressive. It would appear as though the student did not apply their mind when deciding on a treatment rationale and is unable to motivate the reasons for the chosen interventions. The student is unable to identify areas of growth and development due to an inability to accept that there are areas of poor performance requiring growth and development.

2 = Needs improvement

The student is able to accept critique and clinical direction, however somewhat reluctantly. The student lacks the ability to critically analyse their performance. The student whilst being able to identify some areas of growth may not necessarily accept the advice or make an attempt to try and improve. The student may seem sceptical of the advice given by the clinician and consistently defends their point of view.

3=Meets Minimum Standards

The student is able to analyse and reflect on their consultation and treatment of the patient. The student willingly accepts direction and critique from the supervising clinician. The student has the ability to identify areas of personal growth and development. Where differences of opinion arise the student is able to defend their decisions based on evidence based best practices.

4=Exceeds Minimum Standards

The student is a reflective practitioner who makes decisions based on best practice and evidence base. The student accepts direction willingly but is also able to articulately explain/justify therapy treatment regimen using best practice or evidence base. The student knows of alternative treatment protocols. The student makes a concerted attempt to improve and spend time researching other treatment protocols. The student utilises the expertise of the supervising clinician and is open to peer-review.

5 = Far Exceeds Expected Performance

The student has exceeded all objectives and standards in this area by delivering exceptional performance. Functions as a role model, not just meeting, but setting new standards of excellence. Excellent feedback is noted from patients, assistants and assessors where relevant. High levels of expertise and initiatives are demonstrated. Student is recognised by all as contributing constructively to the team through taking a leadership position.

ANNEXURE 4 – UJ CPAI MARK SHEET



DEPARTMENT OF EMERGENCY MEDICAL CARE

CLINICAL ASSESSMENT REPORT

NAME OF STUDENT: _____

STUDENT NO: _____

CLINICAL ASSESSOR: _____

AREA: _____

NAME OF PATIENT: _____

PRF Ref No: _____

DATE OF ASSESSMENT: _____

**DEPARTMENT OF EMERGENCY MEDICAL CARE STUDENT
CLINICAL ASSESSMENT REPORT**

Assessment Area 1: Preparation

Student

Self -Assessment: 1 2 3 4 5

Student's Comments

.....
.....
.....
.....
.....
.....
.....

Assessor: 1 2 3 4 5

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

Assessment Area 2: Professional Presentation

Student

Self -Assessment: 1 2 3 4 5

Student's Comments

.....
.....
.....
.....
.....
.....
.....

Assessor:

1 **2** **3** **4** **5**

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

**Assessment Area 3: History Taking, Communication and
General Patient interaction**

Student

Self –Assessment:

1 **2** **3** **4** **5**

Student's Comments

.....
.....
.....
.....
.....
.....
.....

Assessor:

1 **2** **3** **4** **5**

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

Assessment Area 4: Clinical Assessment and Diagnostics Skills

Student

Self -Assessment: **1** **2** **3** **4** **5**

Student's Comments

.....
.....
.....
.....
.....
.....

Assessor:

1 **2** **3** **4** **5**

Assessor's Comments

.....
.....
.....
.....
.....
.....

Assessment Area 5: Decision-Making relating to clinical interventions

Student

Self -Assessment: **1** **2** **3** **4** **5**

Student's Comments

.....
.....
.....
.....
.....
.....

Assessor: 1 2 3 4 5

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

Assessment Area 6: Clinical, Procedures, Skills and Treatment Interventions

Student Self -Assessment: 1 2 3 4 5

Student's Comments

.....
.....
.....
.....
.....
.....

Assessor: 1 2 3 4 5

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

Assessment Area 7: Clinical Documentation and Record keeping

Student

Self -Assessment:

1

2

3

4

5

Student's Comments

.....
.....
.....
.....
.....
.....

Assessor:

1

2

3

4

5

Assessor's Comments

.....
.....
.....
.....
.....
.....

Assessment Area 8: Handover

Student

Self -Assessment:

1

2

3

4

5

Student's Comments

.....
.....
.....
.....
.....
.....

Assessor:

1

2

3

4

5

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

Assessment Area 9: Ability to reflect on own practice, self-critique and accept clinical direction

Student

Self -Assessment:

1

2

3

4

5

Student's Comments

.....
.....
.....
.....
.....
.....
.....

Assessor:

1

2

3

4

5

Assessor's Comments

.....
.....
.....
.....
.....
.....
.....

DECLARATION BY SUPERVISING CLINICIAN:

I hereby declare that the above assessment was completed during the consultation and treatment of the named patient on this date and all the relevant areas were discussed with the student being assessed after the

clinical session. By attaching my signature to the above I hereby attest to the above.

SIGNATURE

DATE

DECLARATION BY STUDENT:

I hereby declare that the above assessment was thoroughly discussed with me by the Clinician whose name is attached to this document and that I fully understand all areas of my assessment. All my questions have been answered and I have been given advice and guidance towards my development. By attaching my signature to the above document I hereby attest to the above.

SIGNATURE

DATE

ANNEXURE 5– INVITATION TO COMPLETE A QUESTIONNAIRE

10 November 2015

Invitation to Participate in a Masters Research Project

Good day,

My name is Bernard van Tonder. I am currently registered with the Cape Peninsula University of Technology for a Master of Emergency Medical Care, which is a full research Masters. For the qualification, I am conducting research into the validity of the newly developed workplace based clinical performance assessment instrument at the University of Johannesburg. My research topic is **AN INVESTIGATION INTO THE VALIDITY AND RELIABILITY OF AN INSTRUMENT FOR THE ASSESSMENT OF CLINICAL PERFORMANCE DURING WORK INTEGRATED LEARNING OF EMERGENCY MEDICAL CARE STUDENTS AT THE UNIVERSITY OF JOHANNESBURG.**

I would like to invite you to participate in the study. I would like to invite you to complete a survey questionnaire. The questionnaire will ask of you to provide feedback on the UJ CPAI and will take about 10 minutes to complete. Participation in the study is voluntary and the decision not to participate will be respected. You will be allowed to withdraw at any point up until the submission of the questionnaire without incurring any penalties. There will be no direct benefit, monetary or otherwise, from participating in the study. Participation in the study will also not require from you to incur monetary loss.

All of the data gathered will be treated as confidential. Only the supervisors and I will have access to the raw data. The personal information and results will not be distributed in any way.

Please feel free to ask any questions relating to the study to the person explaining it to you. Should you wish to participate, please read the consent form for further instructions.

Yours truly,

Bernard van Tonder

Student: Master of Emergency Medical Care

Cape Peninsula University of Technology

Department of Emergency Medical Sciences

212306189

079 874 2368

**ANNEXURE 6 – SELF-DESIGNED, NON-VALIDATED,
EVIDENCE INFORMED QUESTIONNAIRE**

**AN INVESTIGATION INTO THE VALIDITY AND RELIABILITY OF
AN INSTRUMENT FOR THE ASSESSMENT OF CLINICAL
PERFORMANCE DURING WORK INTEGRATED LEARNING OF
EMERGENCY MEDICAL CARE STUDENTS AT THE UNIVERSITY OF
JOHANNESBURG.**

CONTENT VALIDITY QUESTIONNAIRE

Dear participant,

Thank you for taking the time to complete the questionnaire. The questionnaire will ask of you to provide feedback on the UJ CPAI and will take about 10 minutes to complete. Participation in the study is voluntary and the decision not to participate will be respected. You will be allowed to withdraw at any point up until the submission of the questionnaire without incurring any penalties. There will be no direct benefit, monetary or otherwise, from participating in the study. Participation in the study will also not require from you to incur monetary loss.

I hereby consent to take part in the study.

Please indicate your current clinical role:

External assessor	Lecturer	Clinical facilitator
-------------------	----------	----------------------

Please read the Clinical Practice Assessment Instrument narrative for marking (Supplied). For each assessment area, indicate the extent to which the narrative enables you to accurately assess each assessment area. Use the table below to guide the rating you wish to award.

Rating	Description
1	The narrative does not enable the assessor to assess the area at all.
2	The narrative poorly enables the assessor to assess the area.
3	Neutral, or unsure.
4	The narrative enables the assessor to accurately assess the area.
5	The narrative enables the assessor to assess the area perfectly.

Assessment area 1: Preparation					
Assessment area 2 2: Professional presentation					
Assessment area 3: History taking, communication and general patient interaction					
Assessment area 4: Assessment and diagnostic Skills					
Assessment area 5: Decision making relating to clinical Interventions					
Assessment area 6: Clinical, procedures, skills and treatment interventions					
Assessment area 7: Clinical documentation and record keeping					
Assessment area 8: Handover					
Assessment area 9: Ability to reflect on own practice, self-critique and accept clinical direction					

If you feel that you would like to elaborate on any of the assessment area above, please use the space below. Also, if you have awarded a rating of 3 or less for any of the Assessment area above, please use the space below to elaborate on the reasons why.

Assessment area	Rating awarded	
Reason for rating awarded		

ANNEXURE 7 – INVITATION TO PARTICIPATE IN A FOCUS GROUP DISCUSSION

25 November 2015

Invitation to Participate in a Masters Research Project

Good day,

My name is Bernard van Tonder. I am currently registered with the Cape Peninsula University of Technology for a Master of Emergency Medical Care, which is a full research Masters. For the qualification, I am conducting research into the validity of the newly developed workplace based clinical performance assessment instrument at the University of Johannesburg. My research topic is **AN INVESTIGATION INTO THE VALIDITY AND RELIABILITY OF AN INSTRUMENT FOR THE ASSESSMENT OF CLINICAL PERFORMANCE DURING WORK INTEGRATED LEARNING OF EMERGENCY MEDICAL CARE STUDENTS AT THE UNIVERSITY OF JOHANNESBURG.**

I would like to invite you to participate in a focus group interview for his study. The focus group will take about an hour of your time and will be audio recorded. The recording will be formally transcribed by the researcher. Participation in the study is voluntary and the decision not to participate will be respected. There will be no direct benefit, monetary or otherwise, from participation in the study. There will also not be any direct costs for my account. All of the data gathered will be treated as confidential. Only the supervisors and I will have access to the raw data. The personal information and results will not be distributed in any way.

Please feel free to ask any questions relating to the study to the person explaining it to you. Should you wish to participate, please read the consent form for further instructions.

Yours truly,

Bernard van Tonder

Student: Master of Emergency Medical Care

Cape Peninsula University of Technology

Department of Emergency Medical Sciences

212306189

079 874 2368

ANNEXURE 8 – CONSENT FORM TO PARTICIPATE IN A FOCUS GROUP DISCUSSION

11 December 2015

Consent form to participate in a Master's research study

I,, confirm that the purpose of the study has been explained to me. I have received and read the invitation letter and I understand the contents thereof.

I will be required to participate in a focus group interview, which will be audio recorded. The recording will be sent to a professional company for transcription. Only the researcher and the supervisor will have access to the data, and the data will be kept confidential.

I acknowledge that participation in the interview is voluntary and that I may choose not to take part in the discussion without fear of retribution. I may withdraw from the discussion up until the interview is adjourned.

There will be no direct benefit, monetary or otherwise, from participation in the study. There will also not be any direct costs for my account.

I hereby consent to participate in the study.

Participant's name: _____

Participant's signature: _____

ANNEXURE 9 – CONSENT FORM TO BE AUDIO RECORDED

11 December 2015

Consent to audio record the focus group interview

I,, confirm that the purpose of the study has been explained to me. I have received and read the invitation letter and I understand the contents thereof.

I will be required to participate in a focus group interview. The focus group interview will be audio recorded. The recording will be sent to a professional company for transcription. Only the researcher and the supervisor will have access to the data.

I acknowledge that participation in the interview is voluntary and that I may choose not to take part in the discussion without fear of retribution. I may withdraw from the discussion up until the interview is adjourned.

There will be no direct benefit, monetary or otherwise, from participation in the study. There will also not be any direct costs for my account.

I hereby consent to being audio recorded during the focus group discussion and that the recorded data be used in the study.

Participant's name: _____

Participant's signature: _____

ANNEXURE 10 – PLAN FOR THE FOCUS GROUP DISCUSSION

Introduction – 5 minutes

- **Thank the participants for taking the time to take part**
- **Introduction of the moderator, assistant and participants**
- **Completion of the consent forms**
- **Brief explanation of the agenda for discussion**

Discussion:

Focus area 1 – Construct validity of the CPAI – 10 minutes

Discussion Point 1 (5 minutes)

Some of the feedback from you, the Subject Matter Experts (SMEs) was that was that Assessment Area 1 (Preparation) was only really applicable for the first case.

This may be corrected by doing a number of things. These include that the mark awarded for this area on case one be carried over, or that the narrative be changed so that it can be used on all of the cases being assessed.

- **Are there specific areas in this narrative that could be singled out that is only applicable to the first case?**
- **What changes do you think should be made to assessment area one? Should one of the proposed changes be implemented, or should something else be done?**

Discussion Point 2 (5 minutes)

Another theme from the questionnaire was that it was easy to understand the narrative on when to award marks 1-4 (does not comply until exceeds expectations). It was unclear how students should perform to be awarded a mark of 5.

Do you think that the instrument should include five levels of performance instead of 4?

- **How do you think could the narrative about the levels of performance be changed?**

Focus area 2 – Content validity – 5minutes

Discussion Point 3 – 5 minutes

Assessment Area Three (History taking) was considered by the majority of the SMEs to have poor content validity. Some of the feedback provided referred to including more than one construct being assessed under one area. An example of this is the difference between communicating well, and obtaining a good history.

- **Do you feel that the assessment of more than one construct under one assessment area could negatively impact the content validity of the instrument?**
- **How do you think could the content validity be improved for this assessment area?**

Focus area 3 – Inter Rater reliability – 15 minutes

Discussion Point 6 – 10 minutes

Inter-rater reliability is commonly measured using a correlation coefficient. This means that the scores awarded to a student across all nine assessment areas by two assessors are compared and a correlation coefficient is calculated. A correlation coefficient that is considered to show good correlation or agreement between the assessors is 0.6 or more.

Out of the 52 cases documented, a good correlation was seen 24 times. This means that the data showed that the assessors awarded the same marks less than 50% of the time.

Some authors argue that the most likely reasons for poor correlation coefficients during inter-rater reliability assessments is either because of poor training, or because of the raters interpreting the instrument narrative differently.

- **Did you find the training adequate? Why do you say so?**
- **Did you find that the narrative was written in a way that it was easy to understand exactly how marks should be awarded to students?**

Discussion Point 7 – 5 minutes

An example of an area that was associated with a very different correlation coefficient was assessment area two (Professional Presentation). Area two had a correlation coefficient of -0.37. This means that, one can say with a small degree of certainty that the marks awarded by the raters is likely to differ. This was also the only area with a coefficient of less than 0.4. One argument may be that the assessors understood the narrative very differently and therefore awarded marks very differently.

- **Why do you propose this happened?**

Discussion Point 8 – Support for the use of this instrument - 10 minutes

- **Do you feel that the use of this kind of instrument should be continued in the programme?**
- **What changes do you think could be made to improve the instrument?**

Closing of focus group discussion.
Thank the participants for taking part.
Discussion adjourned.

**NNEXURE 11 – ONE-WAY ANOVA AS WELL AS TWO WAY
RANDOM CHRONBACH'S ALPHA**

Case	One_way_anova	two_way_random
1	0.842	0.779
2	0.060	0.500
3	0.400	0.467
4	0.308	-0.846
5	0.750	0.708
6	0.759	0.646
7	0.891	0.840
8	0.429	0.000
9	1.000	1.000
10	0.375	0.000
11	0.741	0.704
12	0.480	0.502
13	-0.444	-0.889
14	0.895	0.884
15	0.815	0.704
16	0.765	0.652
17	0.711	0.733
18	0.676	0.704
19	0.649	0.685
20	0.833	0.583
21	0.842	0.684
22	0.737	0.684
23	0.769	-1.308
24	0.824	0.735
25	0.389	0.444
26	0.500	0.563
27	0.894	0.859
28	0.750	0.750
29	0.000	0.138
30	0.889	0.556
31	0.906	0.906
32	0.909	0.727
33	-0.375	-0.250
34	0.762	0.444
35	0.313	-0.969
36	0.808	0.692
37	0.800	-0.600

38	0.000	0.118
39	-0.150	-0.300
40	0.600	-0.600
41	-2.571	-2.429
42	0.667	0.704
43	0.231	0.308
44	0.857	0.857
45	0.821	0.821
46	0.000	0.146
47	0.778	0.778
48	0.778	0.778
49	0.774	0.742
50	0.733	0.533
51	1.067	-1.333
52	7.500	-7.000