

**Using a scoring rubric to make
practice-based assessment of pre-
registration nursing students fit for
purpose: A mixed methods study**

Ibraheim Saleh Almalkawi

A thesis submitted in partial fulfilment of the requirements
of London South Bank University for the degree of Doctor
of Philosophy.

March 2019

Word count: 75,747

Abstract

Background

Globally, across all professions, assessing clinical practice has challenged educators for decades. The literature suggests that practice-based assessments are not fit for purpose. Problems relate to subjectivity, inaccuracy, difficulties in understanding the language used in competency statements and distinguishing different levels of competence. In nursing, threats to public safety have been identified when underperforming students enter the nursing register.

Aim

To make practice-based assessment of pre-registration nursing students fit for purpose.

Method

The critical pragmatist epistemology adopted in this study guided a solutions-focused, two-phase mixed method study. The theoretical framework comprised the concepts of assessment *for* learning and authentic assessment.

In Phase one, 100 stakeholders involved in pre-registration nurse education were invited to participate in a two-round Delphi study. In round one, participants (n=48) provided free-text interpretations of three level descriptors for the professional attitude, behaviour and responsibility competency statements in the Pan London Practice Assessment Document. Content analysis was used to develop a draft scoring rubric. In round two, participants (n=51) scored their agreement to the statements in the draft rubric using a five-point Likert scale. A clear stakeholders' consensus (ranging from 86% to 100%) was reached in all categories of the rubric.

In Phase two, the consensus-based scoring rubric was evaluated regarding its effectiveness in enhancing practice-based assessments of student nurses' practice performance. Paired mentors and final year nursing students were recruited (n=51) to use the rubric during a clinical placement. Two, specifically-designed online questionnaires collected quantitative and qualitative data from individual mentors

and students. Descriptive statistics were used for quantitative data. Content analysis was used to analyse open-ended questions.

Results

Completed questionnaires were submitted by 44 nursing students and 39 practice mentors. Merged quantitative and qualitative results revealed that the consensus-based scoring rubric improved mentors' and students' understanding of the language and levels of competency statements, enhanced self-assessment and feedback provision, including enhanced mentor-student mutual understanding and dialogue. Intersubjectivity emerged as an original perspective to further understand practice-based assessment.

Conclusion

The consensus-based scoring rubric made practice-based assessment more fit for purpose.

Recommendations

All disciplines should consider adopting consensus-based scoring rubrics to make practice-based assessments fit for purpose. The psychometric properties of scoring rubrics and their usefulness in various contexts should be researched.

Acknowledgements

This thesis would have not become a reality if it was not for the support and encouragement of others.

My sincere gratitude is for my supervisors Professor Rebecca Jester and Dr Louise Terry. There are no enough words to thank you for your dedication, patience and inspiration. I am most grateful for your expertise in critiquing and challenging my intellectual capacity so delicately. I am very honoured to have you both as my supervisors.

Special thanks for my wife Saeda and my children Hanna, Malak and Yusef. You have sacrificed so much for me and I could not have completed this thesis without your support and encouragement. I am sure you will be delighted that this journey has come to an end.

I would like to acknowledge my fellow PhD students. Thank you Muireann Kelly for mentoring me, Marie Culloty and Jenny Thomas for joining me in the PhD journey with all its ups and downs.

I would like to extend my thanks to my Head of Department, Anthony McGrath, who has been very supportive over the course of this project. Also, I would like to acknowledge my colleagues for their continuous encouragement, especially Graham Carr, I am very grateful to you for freeing me on occasions, so I can “just get this thing done” as you kept saying.

Finally, this study could not have been undertaken without participants. I would like to express my sincere thanks all those who took part and made this study possible. Hopefully, together we have designed a tool that would be useful to our colleagues and students in enhancing assessment in practice.

This thesis is dedicated to my parents, their persistent encouragement of academic achievements kept me going.

List of abbreviations

AMED:	The Allied and Complementary Medicine Database
ANAST	Australian Nursing Standards Assessment Tool
BEME:	Best Evidence Medical and Health Professional Education
BNI:	British Nursing Index
BOS:	Bristol Online Survey
BSc	Bachelor of Science
CINAHL:	Cumulative Index to Nursing and Allied Health Literature
CPPD	Continuous Personal and Professional Development
DARE:	Database of Abstracts of Reviews of Effects
DH:	Department of Health
EBSCO:	Elton B. Stephens Company
EMBASE:	Excerpta Medica database
ENB:	English National Board
ERC:	Education Research Complete
ERIC:	Education Resource information Centre Health Care
EThOS:	Electronic Theses Online Service
HCP:	Health Care Professional
HEE	Higher Education England
HEI:	Higher Education Institution
IR:	Integrative Review
Medline:	Medical Literature Analysis and Retrieval System Online
MMAT:	Mixed Methods Appraisal Tool
NHS:	National Health Service
NHS REC:	National Health Service Research Ethics Committee
NMC:	Nursing and Midwifery Council
OSCE:	Objective Structured Clinical Examination
PEO:	Participants, Educational aspect, Outcomes
PICO:	Population, Intervention, Comparison, Outcome
PLPAD:	Pan London Practice Assessment Document
PRISMA:	Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PsycINFO:	Psychological Information Database
QAA:	Quality Assurance Agency
QCA:	Qualitative Content Analysis
RCN:	Royal College of Nursing
ROI:	Republic of Ireland
SLAiP:	Standards to Support Learning and Assessment in Practice
SPSS:	Statistical Package for the Social Sciences
UK:	United Kingdom
UKCC:	United Kingdom Central Council

TABLE OF CONTENTS

Abstract.....	i
Acknowledgements.....	iii
List of abbreviations.....	v
TABLE OF CONTENTS.....	1
LIST OF TABLES.....	7
LIST OF FIGURES.....	8

CHAPTER 1: INTRODUCTION 9

1.1 Context of the study.....	9
1.1.1 <i>Assessment of students in practice</i>	9
1.1.2 <i>Lack of reliability and validity of the assessment process</i>	11
1.1.3 <i>Failing to protect the public</i>	12
1.1.4 <i>Defining and measuring competence</i>	14
1.2 Rationale and personal motivation for this study	19
1.3 Theoretical perspectives underpinning this study.....	20
1.4 Philosophical perspectives underpinning this study.....	22
1.5 Methodological approach in this study	23
1.6 The research setting.....	24
1.7 Structure of the thesis.....	25
1.8 Chapter summary.....	26

CHAPTER 2: CONTEXTUAL REVIEW 27

2.1 Introduction	27
2.2 Historical background	27
2.2.1 <i>Changes to mentoring</i>	32
2.2.2 <i>Failing to fail</i>	37
2.2.3 <i>The NMC Standards to Support Learning and Assessment in Practice</i>	39
2.3 Scoping exercise to identify factors influencing the quality of mentors' assessment.....	41
2.4 Identifying high impact barriers.....	52

2.5 Chapter summary	54
CHAPTER 3: INTEGRATIVE REVIEW	55
3.1 Introduction	55
3.2 Defining the term 'quality' in this review	55
3.3 The approach used to review the literature	56
3.3.1 Stage 1: Problem identification	58
3.3.2 Stage 2: Literature search.....	59
3.3.3 Stage 3: Data evaluation	65
3.3.4 Stage 4: Data analysis	67
3.3.5 Stage 5: Presentation	86
3.4 Development of the theoretical framework employed in this study	87
3.5 From the theoretical framework to study aims.....	96
3.6 Research aim, questions and objectives	98
3.7 Chapter summary.....	99
CHAPTER 4: ONTOLOGY, EPISTEMOLOGY, METHODOLOGY AND METHODS.....	100
4.1 Introduction	100
4.2 Ontology and epistemology	101
4.3 Methodology.....	106
4.3.1 Mixed methods research	107
4.3.2 Mixed methods validity	111
4.4 Methods	114
4.4.1 Consensus methods	114
4.4.2 The Delphi method	116
4.4.3 Evaluation research	120
4.5 Chapter Summary	123
CHAPTER 5: THE DELPHI	125
5.1 Introduction	125
5.2 Study design	126
5.3 Sample/Sampling	127

5.4 Instruments	132
5.4.1 Round one questionnaire design (qualitative)	133
5.4.2 Round two questionnaire design (quantitative).....	134
5.4.3 Methodological rigour	136
5.5 Procedure	138
5.5.1 Ethical considerations.....	138
5.5.2 Piloting the online questionnaires	140
5.5.3 Inviting and retaining participants	141
5.5.4 Data analysis (round one).....	142
5.5.5 Data analysis (round two)	149
5.6 Results	150
5.6.1 Round one findings (qualitative)	150
5.6.2 Round two results (quantitative).....	152
5.6.3 Findings from free-text responses in round two.....	155
5.7 Chapter Summary	158
 CHAPTER 6: EVALUATION OF THE SCORING RUBRIC	159
6.1 Introduction	159
6.2 Study design	159
6.3 Sample/Sampling	161
6.4 Instruments	164
6.4.1 Questionnaires design	164
6.4.2 Methodological rigour	170
6.5 Procedure	171
6.5.1 Ethical Considerations	172
6.5.2 Piloting the online questionnaires	173
6.5.3 Conducting the study.....	174
6.5.4 Data analysis (closed-ended questions)	174
6.5.5 Data analysis (open-ended questions)	175
6.6 Results	176
6.6.1 Results from the closed-ended questions	178
6.6.2 Findings from the open-ended questions	186

6.7 Chapter Summary	197
CHAPTER 7: DISCUSSION	199
7.1 Synthesis and contextualisation of findings.....	199
7.1.1 <i>Clarity of language and distinction between levels of competence</i>	200
7.1.2 <i>Strengthening the rigour of assessment</i>	205
7.1.3 <i>Enhancement of learning and feedback provision</i>	208
7.1.4 <i>Student-mentor interaction and dialogue</i>	209
7.1.5 <i>Time</i>	211
7.2 Evaluating the overall usefulness of a scoring rubric	213
7.3 Resonance with underpinning theories	217
7.3.1 <i>Assessment for learning</i>	218
7.3.2 <i>Authentic assessment</i>	220
7.3.3 <i>Alternative theoretical consideration - intersubjectivity</i>	221
7.4 Chapter summary	224
CHAPTER 8: CONCLUSION	227
8.1 Introduction	227
8.2 Addressing the research aim.....	227
8.3 Answering the research questions.....	229
8.4 Additional findings	233
8.4.1 <i>Mentor-student interaction and dialogue</i>	233
8.4.2 <i>Time</i>	234
8.5 Strengths and limitations	235
8.6 Conclusions arising from the study	237
8.7 Contributions to knowledge	239
8.7.1 <i>Contributions to the field of practice-based assessment</i>	239
8.7.2 <i>Contributions to the field of nursing philosophy and research</i>	241
8.7.3 <i>Contributions to the field of nursing theory</i>	241
8.8 Recommendations	242
8.8.1 <i>Recommendations for educational institutions</i>	243
8.8.2 <i>Recommendations for healthcare placement providers</i>	244

8.8.3 Recommendations for the NMC and the Pan London Practice Learning Group.....	245
8.8.4 Recommendations for future research.....	247
8.9 Dissemination.....	249
8.10 Thesis conclusion	249
REFERENCES	251
APPENDICES	271
Appendix 1: Copy of the professional values competency statements in the PLPAD	273
Appendix 2: The post-it notes used in the scoping exercise.....	275
Appendix 3: Databases used to conduct the search in the integrative review.	276
Appendix 4: The mixed method assessment tool.....	277
Appendix 5: Round one e-Delphi questionnaire (group one).....	279
Appendix 6: Round two e-Delphi questionnaire.....	282
Appendix 7: University Research Ethical Committee approval for the Delphi study.	293
Appendix 8: University Research Ethical Committee approval for the questionnaire developed after round one as a supplement.....	294
Appendix 9: Invitation email sent to potential participants in round one of the Delphi.	295
Appendix 10: Participants information sheet sent to potential participants in the Delphi	296
Appendix 11: A reminder email to potential participants in round one of the Delphi after two weeks.....	298
Appendix 12: A final reminder email to potential participants in round one of the Delphi after three weeks.....	299
Appendix 13: Invitation email sent to potential participants in round two of the Delphi	300
Appendix 14: A reminder email to potential participants in round two of the Delphi after one week.	301

Appendix 15: A final reminder email to potential participants in round two of the Delphi after two weeks	302
Appendix 16: The first draft version of the scoring rubric after phase one.	303
Appendix 17: Summary of the free-text comments in round two and the actions taken.....	305
Appendix 18: Mentors questionnaire in phase two.	309
Appendix 19: Students questionnaire in phase two.	313
Appendix 20: HEI ethical approval for phase two.....	317
Appendix 21: Permission from Head of Department at the HEI to access students	318
Appendix 22: NHS REC online tool with the decision that the study does not need ethical approval.....	319
Appendix 23: Approvals from individual Trusts to access mentors.....	320
Appendix 24: Indemnity letter from the HEI.	324
Appendix 25: Participants information sheet for phase two (students)	325
Appendix 26: Participants information sheet for phase two (mentors).....	327
Appendix 27: Informed consent form for participating in phase two	329
Appendix 28: A reminder email to participants in phase two after one week.....	330
Appendix 29: A final email reminder to participants in phase two after two weeks	331
Appendix 30: Key messages from the qualitative responses in phase two coded according to their category	332

LIST OF TABLES

Table 2.1: The quality assurance aspects the NMC (2008a) included in the SLAiP. ..	39
Table 2.2: Stakeholders responses.....	42
Table 2.3: Five categories identified in the scoping exercise	43
Table 3.1: The processes category identified in the scoping exercise.....	59
Table 3.2: The PEO model used for framing the research question	60
Table 3.3: PEO synonyms used to expand the essential terms	61
Table 3.4: Inclusion and exclusion criteria	62
Table 3.5: Data extraction sheet to describe the studies included in the integrative review.....	68
Table 4.1: Legitimation framework adapted from Onwuegbuzie and Johnson (2006) and how it was addressed in this research study.	113
Table 4.2: Delphi characteristics (Rowe and Write, 2001) and how they were addressed in this study.	118
Table 4.3: The four dimensions of evaluation (Lincoln and Guba, 1986).	121
Table 5.1: Selection criteria specific to each stakeholder group.	128
Table 5.2: Strategies used in the Delphi to reduce attrition.....	131
Table 5.3: Reliability considerations (Hill and Fowles, 1975) and how they were addressed in this study.	136
Table 5.4: Trustworthiness considerations (Lincoln and Guba,1985) and how they were addressed in this study.	138
Table 5.5: Example of the how data was transferred, prepared and colour coded for similar meaning.....	146
Table 5.6: Units of meaning developed from the colour coded content.	147
Table 5.7: Organising the data according to their units of meaning categories across the level descriptors.....	147
Table 5.8: Draft scoring rubric for statement 1.	148
Table 6.1: Inclusion and exclusion recruitment criteria in phase two.....	162
Table 6.2: The sources of sampling for students and mentors in phase two.....	164
Table 6.3: Items used in the questionnaires and their relation to the research objectives and the findings from the scoping exercise and the IR.	165

LIST OF FIGURES

Figure 3.1: PRISMA flow chart of the literature search	64
Figure 4.1: The definition of core characteristics of mixed methods research 108 (Adapted from Creswell and Plano Clark, 2017).....	108
Figure 4.2: The three core mixed methods designs (Adapted from Creswell and Plano Clark, 2017)	110
Figure 4.3: The mixed methods designs used in the study.....	111
Figure 5.1: Overview of the e-Delphi rounds.....	126
Figure 5.2: Commonly used bench markers for level descriptors.	135
Figure 5.3: Spread of stakeholders' responses in round one	150
Figure 5.4: Spread of stakeholders' responses in round two	152
Figure 5.5: Level of consensus for the 24 items in the questionnaire.....	154
Figure 5.6: Participants' preference for benchmarking terminology	155
Figure 5.7: The final draft scoring rubric.....	156
Figure 6.1: Recognising the required level to achieve	178
Figure 6.3: Help to explain what is expected of students.....	179
Figure 6.4: Distinguishing between different levels of competence	180
Figure 6.5: Understanding the required level to pass	180
Figure 6.6: Directions to understand and assess attitudes and behaviours.....	181
Figure 6.7: Fairness of grading decisions	181
Figure 6.8: Measuring student's performance accurately.....	182
Figure 6.9: Mentors outlook on their own consistency.....	182
Figure 6.10: General outlook on the rubric consistency.....	183
Figure 6.11: Structured approach for identifying learning needs	183
Figure 6.12: Giving accurate feedback to justify the level awarded.....	184
Figure 6.13: Students self-assessment.....	184
Figure 6.14: Reflection on practice and real-life situations.....	185
Figure 6.15: Reflective discussion about professional attitudes and behaviours ...	185
Figure 6.16: Practicality of using the rubric	186

Chapter 1: Introduction

The aim of this chapter is to situate the study context and provide an overview of this thesis. The concept of practice-based assessment across Health Care Professions (HCPs) will be explored¹ before presenting the study setting and rationale.

1.1 Context of the study

1.1.1 Assessment of students in practice

The Oxford dictionary (2018) states that the word ‘assess’ originates from the Latin word ‘assidere’, which means ‘sit by’, implying a close relationship and sharing an experience, where the assessor stays by the learner’s side, observing performance and supporting by identifying learning needs (Gopee, 2015). Another definition is provided by Stuart (2013) who defined ‘assessment’ as a global term that incorporates tests and examinations of coursework, judgement of performance, and any other ways of measuring professional learning. In the context of this study, the term ‘assessment’ denotes the mentors’ judgement of students’ performance in clinical practice placements.

Education programmes for HCPs involve a combination of both theoretical and practice-based elements. Assessing the practice element is an essential part of curricula to evaluate that learners have developed an adequate level of professional capacity (Yorke, 2005). The ultimate purpose of practice-based assessment is to assure the public that graduates are safe and responsible professionals (Trede and Smith, 2012). The term ‘*fitness for practice*’ is commonly used to represent students meeting the requirements and demands of the professional and regulatory bodies and therefore the legitimacy to become registered practitioners.

Within the context of health and social care in the United Kingdom (UK), the term ‘mentorship’ is used to define the process where a qualified practitioner (mentor) is

¹ A more detailed historical and contextual review of assessing nursing students in the UK is provided in chapter two.

assigned to a student (mentee) during a practice placement (Robinson *et al.*, 2012). The term 'mentor' has no unified description and is used interchangeably in different countries and across various HCPs to cover assessor, instructor, preceptor and supervisor (Mead *et al.*, 2011; Chandan and Watts, 2012). However, despite the different descriptions of the role, it ultimately refers to practitioners who provide support and guidance in practice placements.

The Nursing and Midwifery Council (NMC), which is the nursing regulatory body in the UK, was formed when the English National Board (ENB) and the United Kingdom Central Council (UKCC) merged in 2002. The NMC defined a mentor as "A registrant who, following successful completion of an NMC approved mentor preparation programme or comparable preparation (...) has achieved the knowledge, skills and competence required to meet the defined outcomes" (NMC, 2008a, p. 23). The term 'mentor' will be used in this review to describe all clinical practitioners who have undertaken formal preparation for the role enabling them to assess students during practice placements. This entails the responsibilities of facilitating learning *and assessing* students against competencies set by professional bodies.

Pre-registration nurse preparation programmes in the UK went through radical changes in the 1980s when they moved to higher education (Payne *et al.*, 1991; Crotty, 1993). The move saw the nurse tutors' role evolve into university lecturers, with their role in teaching or assessment in practice settings becoming mainly an advisory role (Clifford, 1994; O'Driscoll *et al.*, 2010). The implications of this move had a major impact on registered nurses who needed to fill the gap, signalling a new and formal role for mentors. Their role as mentors had to extend from being only supportive to ultimately having substantial accountability to supervise and assess students in practice and determine whether they are fit for practice (Watson *et al.*, 2002; Nettleton and Bray, 2008). This is commonly known as the 'dual role', which is a unique role only found in the UK and the Republic of Ireland (ROI).

1.1.2 Lack of reliability and validity of the assessment process

The complexity of assessing clinical practice for all HCPs is a long-standing issue and has received much attention for decades. A plethora of studies raise concerns related to the lack of reliability and validity of assessment in practice-based placements including medicine (Dudek *et al.*, 2005; Paisley *et al.*, 2005; Cleland *et al.*, 2008; Sabey and Harris, 2012; Govaerts *et al.*, 2013), occupational therapy (Ilott and Murphy, 1997), and social work (Tanicala and Roberts, 2011; Eno and Kerr, 2013; Rawles, 2013).

Nursing shares the same concerns that mentors' judgements are subjective and do not always accurately reflect students' performance (Ashworth *et al.*, 1999; Brown, 2000; Calman *et al.*, 2002; Pellatt, 2006; Bray and Nettleton, 2007; Yanhua and Watson, 2011). There is reluctance to fail underperforming students in practice settings (Duffy, 2003; Luhanga *et al.*, 2008a; Luhanga *et al.*, 2008b; Larocque and Luhanga, 2013). Internationally, there is evidence of similar concerns reported in Australia (Miller, 2010), Canada (Yonge *et al.*, 2011), Finland (Jokelainen *et al.*, 2013), Ireland (Butler *et al.*, 2011; Bradshaw *et al.*, 2012; Cassidy *et al.*, 2012), New Zealand (Gallagher *et al.*, 2012) and the United States of America (USA) (DeBrew and Lewallen, 2014).

These findings were also supported in the mentorship project that was commissioned by the Royal College of Nursing (RCN, 2015). They undertook a rapid literature review of mentoring models for pre-registration nursing students outside the UK and provided a comparison between UK practice and three models from Amsterdam, Australia and the USA. The review identified clear variation in how students are assessed and who performs the assessment in practice. Differences were also reported in the length of clinical placements, and the dedicated time for mentoring, ranging from 100% in the Amsterdam model to 0% in the UK model.

Although the RCN (2015) review acknowledged that the UK appeared to have the most detailed policy and guidance documents on student nurse mentoring, issues surrounding the difficulties in assessing student nurses in practice are evident

across all different models and practices. Similarly, Helminen *et al.* (2016) examined the international literature related to assessment of clinical practice of student nurses and concluded that inconsistencies in assessment processes vary not only between countries but also between institutions. Helminen *et al.* (2016) also reported that the reliability and validity of assessment tools have rarely been systematically assessed.

1.1.3 Failing to protect the public

The preparation of UK mentors for undertaking an assessor role was inadequate (Robinson *et al.*, 2012), and consequently, concerns emerged about the effectiveness of practice-based assessment, where mentors lacked confidence in carrying out the assessor role, feeling unprepared and unwilling to fail students. 'Failing to fail' became a phenomenon referring to the reluctance of mentors to fail underperforming students (Duffy, 2003; Gainsbury, 2010; Hunt *et al.*, 2012).

Failing to fail underperforming students suggests that mentors are not effective 'gatekeepers'; those who do not meet the professional standards and should have been deemed not to be competent are being allowed to enter the professional register, which places patients at risk (Duffy, 2003; Gainsbury, 2010; Terry, 2013). Duffy's identification of the problem and the threat it poses to public protection led to a considerable volume of empirical literature focussing on failing to fail and the contributing factors. The central emphasis is that assessments need to be defensible to the public in that they correctly distinguish between those who are competent and those who are not, ensuring safety and public protection is not compromised by incompetent nurses (Luhanga *et al.*, 2008a; Gainsbury, 2010; Burden, 2014).

One of the factors contributing to the 'failing to fail' phenomenon, which was frequently reported in the literature relates to the ambiguous terminology of the practice assessment document. The language used in the documents is described as vague and containing too much academic jargon (Brown, 2000; Duffy and Watson, 2001; Neary, 2001; Norman *et al.*, 2002; Dolan, 2003; Duffy, 2003; Scholes, *et al.*,

2004; McCarthy and Murphy, 2008; Miller, 2010; Butler *et al.*, 2011; Fahy *et al.*, 2011; Cassidy *et al.*, 2012; Cassidy *et al.*, 2017; Almalkawi *et al.*, 2018). This was also affirmed in a scoping exercise conducted locally to inform the direction of this study (see section 2.3).

Consequently, mentors experience problems translating and applying assessment outcomes into observable practice activities. Mentors and students have reported spending significant time trying to work out what the NMC competency statements mean rather than assessing the student against them (Neary, 2001; Scholes, *et al.*, 2004). In particular, studies show that mentors find making decisions about the 'softer' aspects of competence such as attitudes and behaviours to be notoriously challenging to define or measure (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016). Thus, when required to justify their decisions regarding students not meeting competency standards, mentors struggle to prove their concerns are justified, hence they feel unprepared and unwilling to fail students (Duffy, 2003; Gainsbury, 2010; Brown *et al.*, 2012).

A link between mentors' reluctance to fail underperforming students and difficulties in identifying the benchmark of what constitutes a pass or a fail is also acknowledged in the literature (Giot, 2000; Neary, 2001; Butler *et al.*, 2011; Heaslip and Scammell, 2012; Cassidy *et al.*, 2017; Almalkawi *et al.*, 2018), as well as the scoping exercise conducted for this study (see section 2.3). In part, mentors struggle to discriminate between different levels of practice relates to the complexity and lack of consensus on what 'competent' means (Watson *et al.*, 2002; Cowan *et al.*, 2007; Gallagher *et al.*, 2012).

There is also evidence that mentors, students and university-based nurse educators have differing views about what is considered an 'acceptable' standard of competence that a student needs to pass (Dolan, 2003; Cassidy, 2009; Cassidy *et al.*, 2017; Almalkawi *et al.*, 2018). Additionally, the ambiguity of what constitutes acceptable levels of competence does not facilitate the provision of tailored and

constructive feedback, which is a vital element in the assessment process (Neary, 2001; Heaslip and Scammell, 2012; Almalkawi *et al.*, 2018).

1.1.4 Defining and measuring competence

A historical perspective on 'competence' was provided by McClelland (1973) who suggested that the term 'competence' originated in the 1930s, and was used as an alternative to academic or general intelligence testing for jobs tended to be manual or not professional. Instead, people could be tested for sets of skills specific to occupations that required them. This suggests that the term 'competence' is viewed as the 'ability to perform', whereby subjecting people to rigorous intelligence testing could be deemed unnecessary.

An alternative explanation was presented by Eraut (1994) who argued that the term 'competence' originated in the nineteenth century when members of associations who saw themselves as skilled practitioners (e.g. architects, engineers, physicians) were concerned about their public status being undermined by those who used the same title without the requisite skills. They perceived their public status threatened by the public's failure to distinguish between those who were competent from those who were not. This led to the introduction of intellectual training and examination to give some degree of assurance about the competence of members of the associations as being 'properly qualified'. Eraut (1994) argued that the frequent reference to examinations as tests of competence, suggests viewing the term 'competence' as a specialised intellectual capability rather than a practical skill, which belongs to the realm of 'trade'.

Examination of the literature reveals that no single definition is accepted and that the term 'competence' is used interchangeably with 'competency' (Eraut, 1994; Watson *et al.*, 2002; Cowan *et al.*, 2005; Yorke, 2005). Eraut (1994) defined 'competence' as a generic quality of a person's overall capability, whereas 'competency' is a narrower concept which refers to specific abilities or episodes, arguably more akin to McClelland's (1973) interpretation. This confusion contributes to the lack of consensus on what the term 'competence' actually means

(Watson *et al.*, 2002; Cowan *et al.*, 2007; Axley, 2008; Cassidy, 2009; Valloze, 2009; Yanhua and Watson, 2011). The ambiguity associated with the term 'competence', is problematic because, according to Watson *et al.* (2002, p. 422) 'competence' refers to "possessing certain qualities by someone without specifying all that they can do in a given set of circumstances". Hence, the confusion about the term 'competence' is historic with a clear distinction between being *properly qualified* and being *able to perform* as a criterion for competence.

Despite the confusion and debate concerning the concept of 'competence' and the difficulty identifying a definition capable of accommodating all the different ways that the term is used (Le Deist and Winterton, 2005; Cowan *et al.*, 2007; Axley, 2008; Valloze, 2009), the term is firmly rooted in professions such as education and healthcare, making the issue of competence and assessment of competence a controversial issue (Barnett, 1994; Watson *et al.*, 2002). In UK nursing, the interest in competences peaked in the 1980s with the move of nurse preparation programmes to Higher Education Institutions (HEIs), replacing state exams with continuous assessment of students' competence in clinical settings (see section 2.2). Since then, it has greatly influenced pre-registration nurse education as the primary driver in measuring effectiveness (Cowan *et al.*, 2007; Gallagher *et al.*, 2012).

Similar to Watson *et al.*'s (2002) description, Yorke (2005) reported the 'competence' denotes effective functioning in the real world, involving not only observable behaviour that can be measured, but also unobservable attributes including attitudes, values and personal dispositions. In the real world of nurse education, as a further complication, 'competence' has become entangled with the concept of 'fitness for practice', referring to the skills required to function effectively and safely in the workplace (NMC, 2002, 2010).

There is also confusion between the two terms: 'competence' and 'performance' (Watson *et al.*, 2002; Cowan *et al.*, 2005). While performance is concerned with being able to do something (Eraut, 2004) or what a person routinely achieves on a

day-to-day basis (Yorke, 2005), consensus is lacking on whether this demonstrates competence or not and whether performance is required in order to demonstrate competence. Thus, competence may represent *potential* to perform, not actual performance (Watson *et al.*, 2002). Govaerts and van der Vleuten (2013) discussed how performance genuinely fluctuates over time and cannot be defined independent of its context. They argued that, to capture competence, it is necessary to focus on aspects that go beyond the technical and context-free aspect of performance. This should include the ability to adapt and flexibly apply and develop knowledge and skills in the face of evolving circumstances. Thus, Govaerts and van der Vleuten's position provides a more coherent explanation of competence in relation to the educational development of student nurses whose practice performance is assessed through a variety of clinical placements where they will increasingly be exposed to evolving and challenging circumstances until they are able to practice independently and safely.

Govaerts and van der Vleuten's (2013) definition of 'competence' that referred to adaptability and flexibly in the face of evolving circumstances creates conflict with the term 'capability'. Capability is defined by Nagarajan and Prabhu (2015) as an integration of knowledge skills and personal qualities used effectively and appropriately in response to varied, familiar and unfamiliar circumstances. Weaver (1982) named the '6 Cs' of capability that an educated person ought to be able to develop as: culture, comprehension, competence, cooperation, creativity and coping. This leads to an inevitable conclusion that capable people are more than competent. Nevertheless, Weaver (1982) defined 'competence' as the application of specialised knowledge such as that of the lawyer, the doctor or the technician. This definition of 'competence' does not echo the definition discussed above that goes beyond the application of skills, making the distinction between 'capability' and 'competence' problematic.

In terms of conceptualising 'competence', two main concepts are reported in the literature: behaviourist and holistic (Watson *et al.*, 2002; Cowan *et al.*, 2005; Gallagher *et al.*, 2012; Garside and Nhemachena, 2013). The behaviourist approach

relies on direct observation of tasks and skills; this approach is common in nurse education referring to students demonstrating the ability to independently implement care. Indeed, the NMC (2002, p. 38) definition of 'competence' reflects this by stating that it is "the skills and ability to practice safely and effectively without the need for direct supervision". This definition has been criticised for focussing on what individuals can do but not what they know, thereby disregarding other attributes to nursing care (Watson *et al.*, 2002; Yorke, 2005). Such attributes to nursing include a wide range of skills that change according to the demands or context and cannot be reduced to a list of competencies (Cowan *et al.*, 2005). Additionally, Watson *et al.* (2002) argued that direct observation and measurement of performance in itself is problematic, and there is no guarantee that generic competencies exist or are sufficient across the board.

The other approach is referred to as 'holistic' since it focuses on general attributes essential for effective performance (McMullan *et al.*, 2003). Here, 'competence' is seen as a cluster of abilities such as motive, personal interest, perceptiveness, maturity, and aspects of personal identity (Cowan *et al.*, 2005; Govaerts and van der Vleuten, 2013), and therefore is regarded as more than the sum of individual competencies. Since 2010, the NMC appear to have advocated a more holistic approach as expressed in the Standards for Pre-registration Nursing Education which bring together several attributes in their definition of competence as "the combination of skills, knowledge, attitudes, values and technical abilities to underpin safe and effective nursing practice and interventions" (NMC, 2010, p. 11). According to Cassidy (2009, p. 41), assessing such holistic and interrelated attributes is necessary for "emotionally intelligent nursing", a concept described by the NMC (2017) as meaning the nurse possesses the values and personal attributes of being caring, empathetic and compassionate.

Despite confusion around the actual definition of the term 'competence', all nurses are required to demonstrate competence against the NMC standards prior to being eligible for registration (Gallagher *et al.*, 2012). This confusion in defining competence, according to Cowan *et al.* (2005), contributes to the variable quality of

mentors' judgement regarding what constitutes an appropriate level to pass or fail a student. Another, as yet unanswered, problem relates to whether there are different levels of competence or only one threshold but differing levels of performance. For example, in the commonly used Benner's (1984) 'novice to expert' framework (discussed in detail in section 3.3.4), 'competent' is placed in the middle, and at that level practitioners are expected to function safely, clearly implying that to be competent, one is not required to be proficient or expert. Eraut (1994, p. 160) described 'competent' as "tolerably good but less than an expert" and Yorke (2005, p. 17) as "good enough". Yorke (2005) explored what 'good enough' really means in practice and identified that a particular problem may lie in the lack of specific criteria against which it can be assessed.

Another fundamental problem relates to discriminating between competent and incompetent students particularly due to assessment tools lacking sensitivity (Norman *et al.*, 2002; Watson *et al.*, 2002; Levett-Jones *et al.*, 2011). Calman *et al.* (2002) identified that, although levels of performance are embedded in the notion of competence, there was no agreed strategy in nursing for describing performance at different levels. The lack of transparent criteria against which students' competence can be judged not only influences the accuracy of understanding and measuring levels of competence, but also how mentors deliver effective and constructive feedback (Fitzgerald *et al.*, 2010; Almalkawi *et al.*, 2018). While there are a variety of assessment taxonomies in existence, ambiguous language and difficulty in identifying performance levels continue to be problematic (Almalkawi *et al.*, 2018). Significant amounts of literature have debated the appropriateness and quality of existing assessment tools (Calman *et al.*, 2002; Norman *et al.*, 2002; Watson *et al.*, 2002; Duffy, 2003; Moore, 2005), yet have produced no resolution to the problem of separating the competent from the not-yet-competent and the incompetent.

Despite the concept of clinical competence in nursing not being universally defined, the term 'competence' adopted in this thesis is the NMC's (2010, p. 145) holistic approach to the definition of competence, as this is what would govern the actual

assessment in practice during the duration of the study. Thus, whenever the term 'competence' is employed, it should be understood to mean: "the combination of skills, knowledge and attitudes, values and technical abilities that underpin safe and effective nursing practice and interventions."

1.2 Rationale and personal motivation for this study

The context summarised thus far illustrates that there is no single uniform method of assessing competence in pre-registration nursing education, resulting in inconsistency and difficulty in interpreting the language used to describe competencies in practice assessment documents. There has been limited exploration or inquiry into how nurses develop a shared understanding of competence, or how different levels of competence can be measured (Terry, 2013; Almalkawi *et al.*, 2018). The available assessment tools largely centre on the use of taxonomies to determine levels of practice, despite criticism of their suitability and appropriateness to assess undergraduate nursing competence (Calman *et al.*, 2002). Therefore, the inconsistency and difficulty of how competency statements are interpreted merits further analysis in order to search for possible solutions that can make practice-based assessment of pre-registration nursing students fit for purpose.

My main motivation for the study stems from my experience as a Senior Lecturer teaching the formal mentor preparation programme since 2009. It was in this role that I became fully aware of the challenges mentors face when assessing nursing students in practice. I recognised very early on the lack of support for mentors in practice and that they are not well equipped or sufficiently prepared in carrying out the assessor role. Yet the language of the regulatory body and the published literature focussed mainly on questioning mentors' reliability and validity of their practice-based assessments, with continuous reminders of their responsibility and accountability as 'gate keepers'.

My aim was pragmatic. I felt passionate about supporting mentors rather than re-establishing the inadequacies of assessment in practice to ensure students' fitness

for practice. As an educator, I also wanted to help students develop competency. This led me to search various ways and ideas that may help both mentors and students through the complexity of assessing students in practice placements.

The next two sections will provide a synopsis of the theoretical and philosophical perspectives that informed this study. These will be addressed in more detail later in the thesis.

1.3 Theoretical perspectives underpinning this study

The integrative review presented in chapter 3, which has been published (Almalkawi *et al.*, 2018), identifies that practice-based assessment does not have one overarching theory. Consequently, a variety of theories that mainly relate to summative assessments such as decision-making or judgement theories seem to be adopted. Thus, the integrative review stresses the need for a suitable theoretical framework to underpin practice-based assessment.

Achieving competence in health care professions such as nursing is directly tied to 'real world' activities, therefore, assessment of competence should involve the *authenticity* of problem-solving in the real world. McDermott *et al.* (2017) protested that in situations where students are assessed on their competence to perform in meaningful activity drawn from a professional context, non-traditional assessment methods that appraise the value of the assessment process itself are required. The design of this study was influenced by two theoretical frameworks that support this notion: 1) assessment *for* learning and 2) authentic assessment (discussed in detail in section 3.4). Both theories are characterised by focussing on application of process and product.

Assessment *for* learning adjusts the balance between formative and summative assessments by recognising the impact assessment has on driving learning, challenging the traditional view that the purpose of assessment is to determine whether a student passes a test. Schuwirth and Van der Vleuten (2011) pointed out that the central purpose of educational curricula is for learning to take place, so

assessment should be aligned with this purpose. Therefore, assessment *for* learning is viewed as a system which directs students learning based on the power of summative assessment in addition to providing feedback (Gibbs and Simpson, 2005).

In their study to evaluate the use of assessment *for* learning in higher education, McDowell *et al.* (2011) reported that when students received guidance and feedback that helped them to understand and meet the requirements without a focus solely on gaining marks, this encouraged them to have responsibility for directing their own learning and gave them a sense that they are personally developing and progressing. This was achieved through understanding the criteria that embodies what it means to do well, shifting the focus towards learning, giving students opportunities to test-out subject knowledge and rehearse relevant skills before these are summatively assessed.

Authentic assessment involves assessing the application of process and product in the real world (Wiggins, 1989; Montgomery, 2002). By taking into account the actual performances the profession wants students to achieve, criteria can be designed to replicate the actual challenges facing students in the real world, helping them to problem-solve, think critically and handle complex tasks (Wiggins, 1989). The expectation is that the achievement would be transferable beyond the immediate learning situation.

What distinguishes assessment *for* learning and authentic assessment is that the criteria are known in advance to focus both the student and the assessor. Central to assessment *for* learning and authentic assessment characteristics is the use of explicit and transparent criteria that describe what the target performance that should be achieved looks like, enabling students and assessors to make formative assessments and continue to modify work in progress to reach their full potential (Wiggins, 1989; Popham, 2001; Montgomery, 2002; Ali, 2013; Haines *et al.*, 2013).

In the literature, authentic assessment is frequently associated with the use of scoring rubrics. A scoring rubric is a matrix with clear and unambiguous description of expectations at different levels of accomplishment that are applied consistently to compare students' performances against the descriptors (Montgomery, 2002; Truemper, 2004; Frentsos, 2013). Scoring rubrics are distinguishable from other grading tools or checklists by providing clearly defined performance criteria written in an easily understood language, making them useful not only as tool to enhance learning, feedback and self-assessment, but also as a tool to judge performance (Allen and Tanner, 2006).

By acknowledging that mentors' lack of consistency (on how competency statements are interpreted or what constitutes an appropriate level of competence), contributes to making practice-based assessment not fit for purpose, this study took a pragmatic approach by searching for alternative ideas to make practice-based assessment of pre-registration nursing students fit for purpose. Guided by the theoretical underpinning of assessment *for* learning and authentic assessment, developing a scoring rubric based on stakeholders' consensus interpretation of level descriptors was considered as a potential solution to the challenges faced in practice-based assessment.

1.4 Philosophical perspectives underpinning this study

The philosophical approach adopted in this study is critical pragmatism (discussed in detail in section 4.2). This contemporary approach accepts the pragmatic principles of practicality and uncertainty about absolute *truth*, and is inspired by Dewey's (1966) idea of knowledge as a critical form of inquiry to find a practical solution that values social action to serve the public good, with emphasis on involvement, participation and openness as a tool to change (Kadlec, 2007).

Dewey saw inquiry as a democratic process that places the focus on practicality and social involvement in deciding what is *truth* and *knowledge*. Dewey explained that to make more sense of the world, ideas are used as instruments of inquiry that require active participation in processes to reconstruct events with alternative

variations, and, testing the practical instrumentality of the proposed change must be carried out experimentally. This dimension of Dewey's pragmatism has been branded with the contemporary name 'critical pragmatism' (Kadlec, 2007). Kadlec (2007) associated critical pragmatism with Habermas's critical social theory, reconstructing Dewey's pragmatic thoughts to appreciate its "critical" features.

The philosophical position of critical pragmatism was a suitable overarching approach for this study because it involved employing an alternative solution (scoring rubric) that ensured social involvement (Delphi consensus method) in deciding what is *truth* and *knowledge* when interpreting the competency statements. This was followed by testing the new approach to evaluate the "practical instrumentality" of the proposed change experimentally.

1.5 Methodological approach in this study

It was recognised that the research objectives could not be achieved through the application of solely qualitative or quantitative methodologies, so a pragmatic strategy of mixing methods was used incorporating two phases. Considering that a collaborative approach in developing tools is essential for its success (Norman *et al.*, 2002; Cassidy *et al.*, 2012), phase one of this study adopted a Delphi consensus method. Delphi consensus (discussed in detail in section 4.4.2) is a method for structuring a group communication process. It consists of two or more rounds of questionnaires administered to a panel. The first questionnaire asks the panel for their opinions on a certain issue or topic in an open-ended manner. These responses are then analysed by the researchers and sent back to the panel in the form of statements or questions. The panel rate or rank the statements or questions within the second questionnaire according to their opinion on the subject. Rounds continue until a consensus is reached on the items as required (Keeney *et al.*, 2011). This method was deemed appropriate to reach stakeholders' agreement on a unified interpretation of level descriptors that would form the blueprint of the scoring rubric.

Phase two of the study evaluated the utility of the consensus-based scoring rubric in terms of its usability, clarity and practicality with mentors and students, as well as its effectiveness as an assessment tool. Evaluation research (discussed in detail in section 4.4.3) is a form of applied research that aims to produce information about implementation, operation and ultimately effectiveness of a programme. The premise is to improve or refine a programme or a system to assess its impact (Lincoln and Guba, 1986). Hence, this method was found suitable for determining the effectiveness and impact of the scoring rubric on the assessment process.

1.6 The research setting

This study was carried out at the School of Health and Social Care in a London University. The School is located at two campuses and delivers the NMC-approved three-year undergraduate pre-registration nursing curriculum on both campuses. The nursing programme is 50% practice-based, as stipulated by the NMC, where students are allocated for their practice placements within a wide range of organisations and institutions across London and the South East.

In their practice placements, students' fitness for practice is assessed against the NMC standards for proficiencies (NMC, 2010)², in the form of a practice assessment document. Since September 2014, the HEIs in the London region have used an NMC-approved single Pan London Practice Assessment Document (PLPAD) to provide uniformity of the practice assessments documents used in clinical practice. This study is based on interpreting the professional attitude, behaviour and responsibility competency statements within the PLPAD (version 2) Adult Nursing Part 3 BSc (2015) presented in Appendix 1. The rationale for focussing on the professional attitude, behaviour and responsibility competency statements will be provided in section 5.4.3.

² The NMC recently published new standards of proficiency (NMC 2018) to be implemented in 2019.

1.7 Structure of the thesis

This thesis is organised into eight chapters. Following this introduction chapter, chapter two presents a contextual review of the background and the current state of practice-based assessment of pre-registration nursing programmes in the UK. The chapter reviews the challenges facing practice-based assessment and the factors that influence the accuracy of mentors' decisions.

Chapter three provides further analysis of the way competency statements in the practice assessment documents are interpreted. An integrative literature review was conducted to systematically synthesise and evaluate empirical and theoretical literature on the challenges mentors face in interpreting and assessing levels of clinical competence in pre-registration nursing. The chapter concludes with the research aim, questions and objectives.

Chapter four discusses the philosophical and methodological approaches that influenced this research. The chapter critically reviews why adopting critical pragmatism was the most suitable ontological and epistemological approach in deciding what is *truth* and *knowledge*. The chapter also explores the methodological approach of mixing research methods to answer different research questions, as well as the rationale for choosing the Delphi consensus and evaluation research to be the most suitable methods.

Chapters five and six describe the methods in detail. This was a two-phase study. Full account of the Delphi (chapter 5) and the evaluation research (chapter 6) methods is presented, with each chapter presenting the associated design, data collection, data analysis and results.

In chapter seven, the discussion integrates the quantitative results and qualitative findings of this research study and critically examines the extent to which the designed scoring rubric enhanced practice-based assessment. The discussion is contextualised by relevant and current literature as well as the theoretical underpinning that informed this research.

Chapter eight concludes the thesis by summarising the study findings and how they addressed the research questions. Methodological considerations and limitations of this research study are discussed. The original contribution to knowledge is highlighted before explaining the implications and making recommendations for HEIs, the Pan London Practice Learning Group and the NMC, taking into consideration the newly published NMC Standards for Student Supervision and Assessment (NMC, 2018a). The chapter concludes by providing recommendations for further research and summarises the national and international dissemination strategies planned for the new knowledge generated by this study.

1.8 Chapter summary

This chapter has provided a summary of the context and rationale for why this research study is both necessary and important. It outlined the structure of the thesis. The next chapter provides a historical and current contextual review of practice-based assessment of pre-registration nursing students.

Chapter 2: Contextual review

2.1 Introduction

The previous chapter discussed the concept of practice-based assessment for HCPs, emphasising the importance of mentors' ability to make accurate decisions about students' competence and ultimately safeguarding the public from incompetent practitioners (Luhanga *et al.*, 2008a). The aim of this chapter is to review a) the historical background to student assessment b) how student nurses are currently assessed during the practice component of their pre-registration nursing programmes in the UK. This is due to the specific nature of mentors' roles in this country.

The chapter will also review the literature concerning the reliability of mentors' decisions when assessing nursing students, including the factors that influence the accuracy of their decision. In line with the tenets of critical pragmatism (discussed in detail in section 4.2), the intention is to identify potential areas that, if modified, could make mentors' assessment of student nurses more fit for purpose.

2.2 Historical background

Historical changes in nurse education have had a significant impact upon assessment of competence in practice. The traditional Nightingale model was based on 'apprenticeship' that had a strong vocational element (Payne, 1997). The phrase 'nurse training' was commonly used to describe the preparation, which was more focussed on practical activity and skills acquisition rather than academic value (Carpenter *et al.*, 2012). Therefore, nursing was seen as subordinate to medicine and described as being ritualistic and based on custom, rather than evidence (White, 1986; Ryan, 1989; Larentzon, 1990; Payne, 1997).

Student nurses were employees, spending the majority of their time in care settings, forming a significant proportion of the workforce for care delivery (Crotty, 1993; RCN, 2007). Learning focussed on disease, and a range of common nursing interventions that were gained through 'hands on' experience under the

supervision of a mentor, defined in 1987 by the ENB as a “wise reliable counsellor” and “trusted adviser” (ENB, 1987, cited in Murray and Staniland, 2010, p. 5). What is important to note in this definition is that mentors did not take part in the practical assessment of students’ suitability for admission to the professional register.

Prior to 1971, nurses were required to demonstrate their clinical competence before examiners appointed by the General Nursing Council in a classroom and under examination conditions (Aggleton *et al.*, 1987). Despite the advantage of consistency, the state exams were subject to criticism for being artificial with concerns raised about transferability to real practice areas (Champers, 1998). Therefore, they eventually became viewed as inaccurate indicators of how well nursing students functioned in unrehearsed and spontaneous practical nursing situations (Yorke, 2005). This led to the introduction of behavioural checklists in the 1970s and continuous practical assessment in the 1980s to provide more natural contexts within which to judge students’ performance (Norman *et al.*, 2002). Continuous assessment in practice brought with it confusion about the role of mentors and their relationship with their students (Champers, 1998). However, the ENB (1989) attempted to clarify the role by re-stating that the role of the mentor was one of counselling and guiding, and as such was distinct from the role of the assessor.

Although the apprenticeship model created nurses who were skilled and experienced at the point of registration (Longley *et al.*, 2007), it came under strong criticism suggesting that students were just a “pair of hands” and their clinical development became secondary to the priorities of the service (Willis, 2012, p. 12). Therefore, learning activities were often inadequate, taking place after the ‘work’ was done (Fulbrook *et al.*, 2000; Glen, 2009; Gillett, 2010), with students being relied upon to provide the labour in a system of constant replacement for the wastage of trained staff (Willis, 2012).³

³ The UK government reintroduced the nurse apprenticeship in 2017 as an alternative route to become a degree-registered nurse. Apprentices train in a range of placement settings, splitting their time between university and their employer with assurances that they will be trained to the same degree of quality and standards as university degrees.

As a result, newly registered nurses felt inadequately prepared to cope with the increasing and changing demands of an evolving health care system which placed more emphasis on health promotion and community care services (UKCC, 1986; Payne, 1991). Consequently, large number of students failed to complete their courses or left the profession upon qualification (RCN, 1985), exacerbating the shortfall in staffing levels required for the clinical areas (Lindop, 1989; Kendrick and Simpson, 1992; Fulbrook *et al.*, 2000).

Subsequently, the Judge report (RCN, 1985) provided detailed proposals for the future of nurse education, and a year later the UKCC adopted most of these proposals when it launched '*Project 2000: A New Preparation for Practice*' (UKCC, 1986). The introduction of Project 2000 brought with it the most radical transformation that challenged the status quo in pre-registration nurse education in the UK (Lathlean, 1989).

Central to implementing Project 2000 was the movement from hospital-based Schools of Nursing to HEIs. The move to HEIs changed the focus from 'nurse training' to 'nurse education', with more focus on critical thinking, problem-solving and delivering evidenced based patient care (Carpenter *et al.*, 2012), and becoming knowledgeable doers and networkers (UKCC, 1986). The purpose of locating nurse education firmly within HEIs was for nursing to be recognised as an academic discipline, with greater emphasis placed on academic achievements that used a wider set of criteria for success than the previous model, hence it was offered at Diploma in Higher Education as the minimum academic award (Fulbrook *et al.*, 2000; Kenny, 2004; Glen, 2009).

It is important to note that following the Department of Health (DH) launch of its '*Modernising Nursing Careers*' project (DH, 2006), and an NMC consultation in 2007, there was a drive to move nurse education to degree-level registration. The rationale was that nursing must become a graduate profession to meet the needs of complex care delivery in an increasingly fast-paced health care system that demands flexible, responsive and highly skilled practitioners (NMC, 2007).

Consequently, since 2010, nursing students in the UK are required to demonstrate knowledge and competence at a minimum of a first-degree level or a postgraduate diploma (NMC, 2010). This move aimed to prepare UK nurses for a global nursing profession that allows them to opt for international mobility. This positions them in line with many other countries throughout the world, fulfilling the 1999 Bologna pledge by all the European Union members (European Commission, 1999).

Another key change with Project 2000 was that students became supernumerary to the workforce rather than part of it, with an intention of developing a nursing programme that intertwined both theory and practice. However, the concept of 'supernumerary' was difficult to grasp and implement in practice. Leonard and Jowett (1990) reported that students still saw themselves as an extra pair of hands when staffing levels were low, being required to do tasks such as bed and tea making. Others were reported being left standing like a statue as some mentors interpreted supernumerary to mean they could only observe (Leonard and Jowett, 1990). Some students too, according to Wilson-Barnett *et al.* (1995), considered that supernumerary status meant that they should observe, listen, and not get involved in providing direct care. This limited the chances for students to integrate and feel part of the team (Leonard and Jowett, 1990).

Serious criticism of Project 2000 developed as studies showed that newly qualified nurses lacked confidence as they no longer had the required skills needed at the point of registration (Luker *et al.*, 1996; McLeod-Clark *et al.*, 1996; May *et al.*, 1997; UKCC, 1999; O'Connor *et al.*, 2001). Therefore, the DH in 1999 launched the 'Making A Difference' strategy, which prompted a review of pre-registration nursing education. Subsequently, in 2001, the UKCC modified the policy in its publication of 'Fitness for Practice and Purpose' altering the balance to 50% practice and 50% theory to emphasise the importance of practice learning and practice-based assessment. The equal balance of theory and practice continues to be maintained in the current and the newly published NMC standards (2010, 2018b).

The move to Project 2000 also saw the demise of the nurse tutors' role, as they evolved into 'university lecturer', which remains current to date. The role of the university lecturers in practice no longer included providing 'hands on' teaching or assessment. Instead they became a 'link lecturer', offering guidance and support to clinical staff and students (Payne *et al.*, 1991; Jowett *et al.*, 1992; Clifford, 1994; Pollards *et al.*, 2007; O'Driscoll *et al.*, 2010). Some were overwhelmed by the pressure of delivering the theoretical component, as well as continuing their own education to pursue academic credibility (Elkan and Robinson, 1995). Unlike the medical profession model that allowed doctors to move effortlessly between clinical and academic roles, nurses who wished to progress into academia had to take an academic career pathway and compromise their practical clinical credibility and career (Taylor *et al.*, 2010).

Accordingly, lecturers' responsibilities for the practice element were a source of confusion. Despite the UKCC (1986) and the DH (1989) stating that link lecturers must be clinically credible, they did not provide details of how this could be achieved. For example, the Project 2000 document stated that teaching in the practice settings should not be left to clinical staff (UKCC, 1986), but it failed to provide clarification or examples of what this would entail (Elkan and Robinson, 1995). A study by Crotty (1993) illustrated how link lecturers were 'baffled' by the notion that they could be clinically credible without providing direct care.

Hunt (2014) found in her study that link lecturers struggled between identifying themselves as being part of 'the university' with its corporate goals or seeing themselves primarily as nurses. She reported that lecturers with strong links to practice placements were keen to maintain their professional integrity before meeting the requirements of their employer. This may reflect the opinion of link lecturers in Clifford's (1993) study who expressed that they should have a role to play in assessing the clinical practice element of student nurses. In contrast, Hunt (2014) found that some lecturers absorbed the university culture and seemed to be detached from the reality of how mentors had to function. Consequently, they tended to underestimate the difficulty of working alongside a student for most of

the day, with some lecturers being impatient with mentors who struggled with challenging students (Hunt, 2014).

2.2.1 Changes to mentoring

The role of mentor as assessor

Following the considerable reduction of nurse lecturers' clinical involvement, their role was passed onto nurses in the clinical settings who had not been fully informed about the new role or prepared for the programme. They also had very little guidance on how to act as mentors responsible for supporting, supervising and assessing students in practice settings and making judgements regarding their competence (Andrews and Wallis, 1999; Watson *et al.*, 2002; Nettleton and Bray, 2008; Robinson *et al.*, 2012).

Bestowing the assessment element on clinically-based nurses marked the formal introduction of mentorship as it stands currently, which is unique not only against other professions but also compared to nursing education in other countries (Fulton *et al.*, 2007). The continuous assessment model involved three formal meetings between the students and their mentors: an initial interview to agree the learning objectives of the placement; a mid-point interview to formatively review progress and address any developmental concerns, and a summative final interview to make judgement as to whether the student should pass or fail the clinical placement.

Yet, the ENB did not include formal supervision or assessment when the role of the mentor was redefined in 1989 as 'someone selected by the student to assist, befriend, guide, advice and counsel' (ENB, 1989), implying that the role of mentor and that of the assessor are separate. It was not until 2001 when the 'dual role' of mentors appeared in documentation from the ENB and the DH defining the role of a mentor as "a nurse, midwife or health visitor who facilitates learning and supervises and assesses students in the practice setting" (ENB/DH, 2001, P. 6). This definition was also endorsed by the NMC when they replaced both the UKCC and ENB (NMC, 2004a, 2010). The dual role, where mentors judge a student at the same time as being their counsellor and friend, was widely criticised for creating a conflict of

interest, and compromising mentors' objectivity when assessing performance (Neary, 1997; Bray and Nettleton, 2007). Additionally, the study by Bray and Nettleton (2007) reported confusion and lack of clarity about the dual role as mentors identified their role to include teaching, supporting, listening and passing on clinical knowledge, but none of the mentors interviewed identified assessment as an aspect of the mentor role.

In 2018, the NMC launched the new Standards for Student Supervision and Assessment that cover learning and assessment in practice to be implemented in 2019. The new standards provide radical and ambitious changes to the way HEIs and their practice partners educate student nurses. This includes introducing three roles comprising practice supervisor, practice assessor and academic assessor to work in a tripartite model to assess and confirm students' practice and academic achievement. This abandons the dual role where mentors are simultaneously the facilitator and the assessor. These two roles will be separated into a practice supervisor and a practice assessor and cannot be performed by the same person.

However, the practice assessor role will continue to be performed by registered nurses who will "conduct assessments to confirm student achievement of proficiencies and programme outcomes for practice learning" (NMC, 2018, p. 9). Therefore, considering that practice assessors may spend less clinical time with students, be responsible for assessing more students and reliant on the feedback of practice supervisors to inform their decisions, it is difficult to envisage how the new standards will make a positive change to the current problems facing mentors in ensuring that practice-based assessment is fit for purpose.

Preparation for the mentor role

The speed with which Project 2000 was implemented, particularly introducing the formal system of mentoring with regards to assessment responsibilities, prevented HEIs from preparing thoroughly (Payne *et al.*, 1991). The low staffing levels due to the workforce changes to student nurses' status put extra pressure on service demands, therefore finding time for supervision was a struggle (Elkan and Robinson,

1995). Furthermore, with the ward sisters/charge nurses' role evolving to ward managers, taking on more managerial and administrative work that limited their presence in the clinical areas, more junior nurses became responsible for practice assessments (Phillips *et al.*, 1996). There were reports that some mentors had only been qualified as a registered nurse for less than six months (Watson, 1999). Research has shown that newly qualified nurses experience a stressful transition process for the first 6-12 months (Halpin *et al.*, 2017), supporting the NMC's (2008) criteria for nurses intending to take on the role of mentor that they must be registered for at least one year. This requirement which, in effect, means allowing time for new nurses to consolidate their own practice, does not seem to be featured in the new NMC standards.

The ongoing concerns that mentors lacked preparation and needed support themselves to fulfil this role was reflected in the guidelines for preparation for teachers and mentors document (ENB/DH, 2001), which led the HEIs to introduce the role of practice educators, with a remit including supporting students and mentors and working with them jointly (Robinson *et al.*, 2012). The practice educator role was poorly regulated and defined; hence, it has various titles including: clinical placement co-ordinator, practice placement facilitator, practice development facilitator, clinical practice facilitator and practice education facilitator⁴. The role is very demanding and with practice educator to student ratios of up to one to fifty (RCN, 2015), finding time to work with students proves very challenging.

In some areas, a lecturer practitioner role was created to further facilitate the integration of theory and practice. Williamson and Webb (2001) reviewed the lecturer practitioner posts and reported that these roles helped bridge the theory-practice gap for both HEIs and clinical practitioners. However, students reported little experience of support from lecturer practitioners, and the insufficient time spent with students was attributed to the conflict in competing demands of

⁴ For consistency, the term 'practice educator' is used throughout the study.

balancing the clinical and educational aspects of the role to meet the expectations of different organisations (Williamson and Webb, 2001). Thus, studies reported that students rated practice educators as more likely to be supportive than lecturer practitioners during the placements (Clarke *et al.*, 2003).

What constituted adequate mentor preparation for undertaking the role was not well addressed, and there was no regulatory requirement for mentors to have dedicated training and education on the role and its responsibilities (Andrews and Wallis, 1999). Hence, mentorship preparation varied with some organisations offering local short workshops while others failed to recognise mentorship as a priority for investment, making it difficult for mentors to undertake training and updates (Jones, 2005). Clutterbuck (2004) added that the need to support organisational performance targets related to ensuring that increasing numbers of students were each allocated a mentor, with little thought given to the nature and type of mentorship scheme adopted.

Even when the ENB introduced the teaching and assessing in the clinical practice course (the ENB 998 course) that enabled formal recognition as a mentor, there were reports that the content and effectiveness of the ENB 998 course varied and was often inadequate (Neary *et al.*, 1994; Robinson *et al.*, 2012). It was claimed that, in reality, most mentors learnt on the job (Wilson-Barnett *et al.*, 1995).

In 2008, the NMC launched a developmental framework for the mentorship preparation programme as part of the standards to support learning and assessment in practice (reviewed in section 2.4.3). However, the literature continues to report problems with the mentorship system not being fit for purpose, particularly mentors expressing doubt about their ability to rigorously and fairly assess students (Cassidy, 2009; Gainsbury, 2010; Javis and Tilki, 2011; Gallagher *et al.*, 2012; Hunt *et al.*, 2012; Gopee, 2015). The new NMC standards do not specify a mentor preparation programme like the current standards, they simply stipulate that HEIs and their practice partners must ensure that practice assessors undertake preparation or evidence prior learning and experience to enable them to conduct

objective and evidence-based assessments and provide constructive feedback (NMC, 2018). Lack of clear requirements to prepare practice assessors may reintroduce the problems associated with having various preparation programmes as discussed in the previous section.

Quality assuring practice placements and assessment

In contrast to the sometimes ill-prepared, and inexperienced mentors supporting the practical component of the course (Elkan and Robinson, 1995; Robinson *et al.*, 2012), nurse lecturers, who support and assess the theoretical element, are experienced practitioners who hold a post graduate teaching qualification (NMC, 2008a). Additionally, unlike mentors who had the role imposed on them (Robinson *et al.*, 2012), nurse lecturers have chosen teaching and assessing as a career pathway, suggesting that they are more enthusiastic about the role.

Yorke (2005) identified that practice placements fell largely outside the control of the HEIs, and hence could not be subjected to the rigorous quality assurance procedures operated in respect of the academic component. A report commissioned by the National Health Service (NHS) in London described how assessing students in practice, where decisions are made by a single person, was regarded as less robust than assessment in higher education, where there is an established process for appraising the standards of students' work through marking, moderation, external examining and assessment boards (Robinson *et al.*, 2012). Although HEIs review the practice assessment documents, this is usually limited to retrospectively ensuring that the process of completing the document has been followed, rarely challenging the final decision made by the mentor to pass or fail the student.

Consequently, concerns have emerged about the effectiveness of practice-based assessment, with many studies questioning mentors' reliability and validity of practice-based assessment, declaring that their judgements are subjective and do not always accurately reflect students' performance (Ashworth *et al.*, 1999; Brown,

2000; Calman *et al.*, 2002; Pellatt, 2006; Bray and Nettleton, 2007; Yanhua and Watson, 2011).

2.2.2 Failing to fail

The phrase 'failure to fail' was initially drawn to the profession's attention by Lankshear (1990) describing how it was difficult to fail students in practice. Since then, the phrase became a phenomenon referring to the reluctance of mentors to fail underperforming students (Duffy, 2003; Gainsbury, 2010; Hunt *et al.*, 2012; Duffy, 2013). Duffy's (2003) influential study was commissioned by the NMC to investigate why some mentors were passing students they thought should have failed. Her findings confirmed that nursing students were passing clinical assessments even when there were doubts about their clinical performance. The mentors identified that failing a student was a difficult thing to do and that personal, emotional, as well as, practical issues influenced their judgements. Subsequently, students were often given the benefit of the doubt and so they progressed through the system without having demonstrated the required level of competence.

In 2010, despite the NMC attempts to strengthen the mentor role (discussed in the next section), unpublished research from the University of Hertfordshire showed that mentors were still failing to fail underperforming students, seven years after the issue was first raised by Duffy in 2003 (Lawson, 2010). The same year, a survey of nearly 2000 mentors found that 37% had passed students even though they had concerns about their competence or attitude (Gainsbury, 2010). These findings reflect the magnitude and persistence of the problems associated with practice-based assessment not being fit for purpose, allowing students who do not meet required standards to gain professional registration.

The problem of failing to fail nursing students in practice-based assessments is international, with literature from Australia (Terry, 2013), Canada (Larocque *et al.*, 2013), Ireland (McCarthy and Murphy, 2008; Cassidy *et al.*, 2012), Malaysia (Enrico and Chapman, 2011), Scandinavia (Jokelainen *et al.*, 2013), Singapore (Jinks and

Harron-Iqbal, 2002) and the United States (Cangelosi *et al.*, 2009) reporting the reluctance of mentors (or their equivalents) to fail students. It is also important to note that feeling unprepared or unwilling to report students' failing performance is not unique to nursing as there are similar reports in other professions. This includes medicine (Tonesk and Buchanan, 1987; Cohen *et al.*, 1993; Speer *et al.*, 1996; Hatala and Norman, 1999; Cleland *et al.*, 2008), dentistry (Licari and Chambers, 2008; Willis, 2009), occupational therapy (Whiteford, 2007) and social work (Sharp and Danby, 2000; Finch, 2012; Finch and Taylor, 2013).

Failing to fail appears to have contributed to the low failure rates in practice compared to academia (Calman *et al.*, 2002; Jevis and Tilki, 2011; Hunt *et al.*, 2012). Hunt *et al.* (2012) recognised that there was no clear picture of pass and fail rates in practice assessment and therefore conducted a national survey to compare failure rates between theory and practice. The study found that failing the theoretical components outstripped practice by a ratio of five to one; the data also provided evidence that in some HEIs, students never failed the practice component.

Hunt *et al.* (2012) stated that it could be argued that support for students in practice settings was so effective that the majority were able to achieve the required level of competence. However, this is contradicted by the strong evidence discussed so far in this review that mentors avoid failing underperforming students in practice assessments (Duffy, 2003; Gainsbury, 2010; Lawson, 2010), and that the practice element of the programme is not viewed as important as the theoretical element (Yorke, 2005). The concerns about reliability and validity of mentors' judgements mean that some mentors are not effective 'gatekeepers', and that it may be possible for unsafe and incompetent students to enter the professional register, thereby compromising patients' safety (Duffy, 2003; Luhaga *et al.*, 2008a; Luhaga *et al.*, 2008b; Gaisbury, 2010; Larocque and Luhanga, 2013; Gopee, 2015).

Acknowledging that failing to fail could result in unsafe students joining the professional register and deemed fit for practice, coupled with the evidence that the level of fitness for practice concerns referred to the NMC continue to increase

(NMC 2018c), no literature was found that looked at the relationship between failing to fail students and the increasing referral rates to the NMC. Although not a specific focus of this research study, establishing if such a relationship exists and whether the strategies to strengthen mentoring introduced by the NMC's (2008) standards to support learning and assessment in practice (see next section) succeeded in reducing fitness for practice referrals would be an area for future research.

2.2.3 The NMC Standards to Support Learning and Assessment in Practice

The UKCC set the 'Standards for the Preparation of Teachers of Nurses, Midwives and Specialist community public health nurses' in 1999, which were adopted and republished by the NMC in 2002. Integrating recommendations from Duffy's (2003) study, the learning and assessment in practice consultation (NMC, 2004b), and review of fitness for practice at the point of registration consultation (NMC, 2005), the NMC introduced the Standards to Support Learning and Assessment in Practice (SLAiP) in 2006 which were implemented in 2007 (NMC, 2008a).

While it was recognised in the NMC Code that all registrants have a duty to support students to help them develop their professional competence (NMC, 2008b, 2015, 2018d), the standards introduced quality assurance strategies to promote mentors' awareness of their accountability to public protection through failing unsafe students. Table 2.1 provides a summary of the quality assurance aspects the NMC (2008a) included in the SLAiP.

Table 2.1: The quality assurance aspects the NMC (2008a) included in the SLAiP.	
Criteria to be a mentor	Mentors must have held their registration for at least a year and undertaken an approved mentorship preparation course.
Mentorship preparation programmes	Emphasis on the skills needed to assess competence and must be 10 days spread over 3 months split between HEI and practice. 5 days should be protected time.
Mentor supervisors	Mentor supervisors must sign-off the practical component and have an up to date entry on the register of mentors but do not have to be a sign-off mentor.
Local register of mentors	The mentorship qualification is recorded locally within Trusts rather than on the professional register.
Due regard	Only a registered nurse may sign-off a nursing student at the end of a placement and must be on the same part of the register.

Cont. Table 2.1: The quality assurance aspects the NMC (2008a) included in the SLAiP.	
Delivery in practice	Mentors must not have more than 3 students at any one time. Students should spend 40% of their placement working under the supervision of a named mentor.
Annual updating	Mentors are expected to attend annual updates to ensure that they are informed of issues and changes in pre-registration nurse education.
Triennial review	A check on each mentor must be made every three years to ensure that they have undertaken annual updates, mentored at least 2 students during the 3 years and have kept a portfolio of evidence of mentoring activities.
Sign-off mentors	Have to meet a specific criterion that entails them to be assigned to the student in their last placement signing them off as fit for practice. Sign-off mentors must have an hour a week of protected time with each student.

The Standards instructed institutions providing placements for students to regulate and monitor mentors by initiating and maintaining a 'live mentor register'. To stay on the register, mentors are required to have yearly updates and a portfolio to illustrate reflection on their learning. The standards also introduced the 'sign-off mentor' as a mentor who has met additional NMC criteria. While all mentors may assess individual competencies, only sign-off mentors are designated as being able to sign-off proficiencies at the end of a programme to ensure that students have reached the required standard of proficiency for entry to the NMC register.

However, Robinson *et al.* (2012) criticised the sign-off mentor role, claiming that mentors gave students the benefit of the doubt on the knowledge that a sign-off mentor will be assessing them in the final placement, and therefore contributed to the phenomenon of failing to fail. As discussed earlier, Gainsbury (2010) and Lawson (2010) established that the issue of failing to fail, reported by Duffy (2003), continued to prevail despite the quality assurances effort in the form of the SLAiP to strengthen the role of mentors.

At the time, the NMC response to these reports was to remind HEIs and their partner placement providers of their responsibilities in ensuring that all NMC standards for all approved programmes are met in full with regards to assessment, and to ensure that students finishing pre-registration programmes are able to practise safely and effectively (Weir-Hughes, 2010, cited in Gainsbury, 2010). It is worth noting that the new NMC standards (NMC, 2018a) do not incorporate any of the strategies to strengthen mentoring described in Table 2.1, suggesting that they

are no longer deemed a requirement. It is not clear at this stage whether practice placement providers across the UK will continue to use them, or if there will be a transitional period for pre-registration courses starting prior to implementing the new standards.

In summary, the radical changes to pre-registration nursing programmes in the UK saw the introduction of the formal role of mentors to determine student nurses' fitness for practice. Concerns about the reliability and validity of mentors' judgements means that some mentors fail to fail students who do not meet required standards, compromising public safety. The approach in this study was pragmatic, aiming to identify possible solutions that may help both mentors and students. The intention was to identify factors contributing to the process of failing to fail that have a big impact on practice-based assessment and examine if modifying such factors improve the outcome. To identify and categorise the real problematic areas, a scoping exercise with stakeholders involved in pre-registration nurse education was conducted locally, which will be discussed next.

2.3 Scoping exercise to identify factors influencing the quality of mentors' assessment.

The need to support mentors to do their role is well documented in the literature (Wilson-Barnett *et al.*, 1995; Duffy, 2003; Rutkowski, 2007; Nettleton and Bray, 2008; Webb and Shakespeare, 2008). There is, however, no clarity about what or how the support should be delivered. Although promoting awareness of roles, responsibilities and accountability to encourage mentors to rigorously scrutinise students in practice is central to ensuring the public is protected from unsafe practitioners, examining the root of the problems (with the intention of acting on the barriers that may hinder mentors' assessment of nursing students in practice) is the focus of this research study.

Therefore, in line with the critical pragmatist philosophy (explained in the next chapter), this contextual review included the conduction of a scoping exercise that

involved local stakeholders representing key roles in the provision of pre-registration nurse education and mentorship across one HEI and two NHS Trusts comprising five hospitals (Table 2.2).

Table 2.2: Stakeholders responses	
Stakeholders who took part in the scoping exercise	Number
Student nurses	45
Mentors (From 2 NHS Trusts)	19
Lecturers (module leaders and link lecturers)	8
Pre-registration Course Director	3
Pre-registration Programme Director	1
Practice Educators (From 2 NHS Trusts)	4
NHS Trust Head of Nursing	1

The participants (n=81) were asked to comment on what they considered were the factors influencing the quality of mentors' assessments in clinical practice. The findings from the scoping exercise were related back to published literature to identify the level of similarity or difference found in the local area compared with the national and international picture. The purpose of the scoping exercise was to identify barriers that were perceived to have the highest impact on the quality of practice-based assessment to inform the literature search strategy for the integrative literature review presented in chapter 3.

Apart from the students, who were accessed while attending their taught component at the HEI, all other participants were accessed at meetings related to mentorship. To maintain anonymity and to eliminate any individuals from dominating or influencing responses, participants were given a small piece of paper stating the following question: "what are the factors influencing the quality of mentors' assessment?" A post-it note was also provided on which to write their comments (see Appendix 2). Since the aim of the exercise was only to seek written comments, formal ethical approval was not required. However, a full explanation was provided emphasising that participation is voluntary and completion implied consent.

Considering the amount of data was small and the purpose was to identify areas of difficulties raised by the stakeholders rather than interpreting emerging themes,

content analysis was deemed an appropriate approach to categorise the key factors while staying faithful to the comments made by participants (Finfgeld-Connett, 2014). The process started by being immersed in the data to become familiar with it in order to identify words or phrases that could be used to generate units of analysis. Statements with the same meaning or similar content were then classified into five categories: organisational pressures; processes; relationships and emotions; mentors' aptitude, and decision-making (Table 2.3).

In the organisation and reporting phases of the content analysis, the five categories were structured in accordance with the frequency of them being cited by the stakeholders. Starting with the most frequently cited issues, the assumption was that a higher impact could be achieved by altering or redesigning the most repeatedly stated barriers. The following review of the literature is structured in accordance with the frequency of each category being cited by the stakeholders.

Table 2.3: Five categories identified in the scoping exercise	
1. Organisational pressures	Lack of time/extra time when failing a student. Staffing level/number of students. Patient dependency/acuity/clinical pressure/priority of mentors.
2. Processes	Understanding/familiarity with the assessment documents. Understanding of competency/criteria. Understand what need to be demonstrated to be worthy of a pass. Identifying levels of performance. Accurate/constructive feedback. Inconsistency between assessors.
3. Relationships and emotions	Unwilling to appraise student- worry about their feelings. Reluctance to write negative remarks. Emotional pressure (guilt/fear). Response to feedback. Friendly/unfriendly relationship/like or dislike the student.
4. Mentors' aptitude	Aptitude for mentoring. Understanding the role of the mentor. Conflicting role/dual role. Relevant training /update to perform the task. Reluctance to engage in the process. Mentors' own knowledge and experience.
5. Decision-making	Confidence to fail students. Uncertainty of programme or students' needs. Benefit of the doubt for borderline student. Assessing attitudes. First impression/intuition.

Organisational pressures

Time was the most frequently reported barrier in the scoping exercise. Clinical workload, staff shortages and the conflicting demands of providing patient care created significant barriers, preventing mentors from having sufficient time to prepare, supervise and assess students effectively (Atkins and Williams, 1995; Pulsford *et al.*, 2002; Watson *et al.*, 2002; Webb and Shakespeare, 2008; Veerahmah, 2012). This was made worse by the increased demands for placement capacity in already oversubscribed clinical areas, posing a real challenge to find sufficient and appropriately trained mentors who met the NMC standards, to support learning and assessment of practice (Nettleton and Bray, 2008).

Limited observation of the student's performance due to lack of time may compromise mentors' judgement, undermining the whole assessment process (Webb and Shakespeare, 2008) especially when dealing with a failing student (Wilkes, 2006). Studies reported mentors being unable to complete assessments on time (Duffy, 2003), with evidence of dedicated mentors using their own time to complete the paperwork either out of hours or in their own homes (Atkins and Williams, 1995; Dolan, 2003).

Despite the recognition by the DH (1999), the NMC (2008a, 2018e) and the RCN (2015) that mentors should have protected time to function adequately in their role (something advocated for in the literature for decades: Atkins and Williams, 1995; Nettleton and Bray, 2008; Veerahmah, 2012), it is very difficult to envisage that mentors being given protected time could happen. Even meeting the NMC standard of sign-off mentors having one hour a week protected time with their students has not materialised, with the sign-off mentors reporting using their own time to meet with the students (Robinson *et al.*, 2012). Lack of time, according to the RCN (2015) is a key issue as to why mentors cannot fulfil their duties, and why there is a reluctance to become a mentor when other registrants see the extra demands required of this role. Hence, as Hunt (2016) found in her study, the 'gate-keeping' of the nursing profession relies on mentors' willingness to give up their personal time.

The new NMC Standards for Student Supervision and Assessment continue to stress that practice supervisors and assessors should have sufficient opportunities to “periodically observe the student across environments in order to inform decisions for assessment and progression” (NMC, 2018a, p. 9). However, the RCN confirms that protected time for mentors to fulfil the role has not been addressed by those who have the power to effect change (RCN, 2015). There have also been unimplemented recommendations for building a financial model into organisational planning with estimates for the mentoring time (Pulsford *et al.*, 2002; Webb and Shakespeare, 2008).

Since August 2017, nursing students in England no longer receive a NHS bursary. Instead they are paying university fees with access to the same student loan system as other university students. It is envisaged that self-funding students will have a more critical view and be more assertive about their learning, mentoring and assessment. The impact of this on placements is yet to be known, but it is likely that the self-funding student will be demanding more time with their mentors to develop clinical competence or may become more confrontational about being assessed as incompetent. The financial arrangements and how practice placement providers are paid (in light of nursing students paying), including the impact on students’ expectations regarding teaching, learning and assessment in clinical practice, is a debate that still needs to be had.

Processes

Understanding and familiarity with the assessment document, with particular reference to the ambiguous terminology being difficult to grasp, was the second most frequently identified barrier in the scoping exercise after the issues related to lack of time. A considerable amount of literature confirms participants’ views that the language used is vague and packed with academic jargon (Brown, 2000; Neary, 2000; Duffy and Watson, 2001; Norman *et al.*, 2002; Dolan, 2003; Duffy, 2003; Scholes *et al.*, 2004; McCarthy and Murphy, 2008; Miller, 2010; Butler *et al.*, 2011; Fahy *et al.*, 2011; Cassidy *et al.*, 2012).

Consequently, mentors experienced problems translating and applying assessment outcomes into observable practice activities, which led to problems in knowing how to accurately assess learning and assign a grade or how to use the tools consistently and effectively. Hence, they had no confidence in practice documents (Norman *et al.*, 2002; Scholes *et al.*, 2004). Both mentors and students have reported spending significant time trying to work out what the competency statements mean rather than assessing the student against them (Neary, 2000; Scholes *et al.*, 2004) and have consequently reported ending up negotiating their own objectives and learning outcomes. Thus, when it comes to justifying their decisions, mentors struggled to prove their concerns were valid (Duffy, 2003; Gainsbury, 2010; Brown *et al.*, 2012).

Difficulties in discriminating between different levels of practice were also frequently reported in the scoping exercise. The literature acknowledges that mentors struggle to identify the benchmark of what constitutes a pass or a fail (Norman *et al.*, 2002; Cowen *et al.*, 2005). This was most noticeable when dealing with borderline students (Duffy, 2003). As discussed earlier (see section 1.1.4), in part, this is directly related to the complexity and lack of consensus on what 'competent' really means, but there is also evidence that mentors have differing views about what is considered as the 'acceptable' standard of competence a student needs to meet in order to pass (Cassidy, 2009).

Neary (2001) evaluated the various tools or systems used to grade practice performance and found that it was not clear that all those involved were fully aware of the meaning or value of the particular grading scheme used. This may be explained by the fact that the current grading tools provide generic descriptors that lack specificity so remain open to interpretations. The lack of clear criteria against which students' performance can be judged, not only influences the accuracy of completing students' documents (Fitzgerald *et al.*, 2011), but also how mentors deliver effective and constructive feedback (Neary, 2001).

Relationships and emotions

Several comments in the scoping exercise referred to the emotional side of the mentor-student relationship. On reviewing the literature, it appeared that most of the issues stem from the conflict between the supportive and facilitative aspects of the role (Robinson *et al.*, 2012). Mentors are encouraged to build supportive and nurturing relationships for the pastoral aspect of the role, but this could also be harmful as getting to know the student over the period of the practice placement may compromise the objectivity of the summative assessment (Duffy, 2003). Hence, developing and maintaining a good relationship that balanced being a friend and an assessor is seen as emotionally labour intensive (Webb and Shakespeare, 2008).

There are also reports that mentors found giving honest feedback about negative aspects of practice a stressful experience (Dolan, 2003). Some would refrain from sharing their opinion with the students to avoid upsetting them (Duffy, 2003; DeBrew and Lewallen, 2014; Helminen *et al.*, 2016) or for fear of breakdown in the relationship (Fotheringham, 2011). Fear of reprisal or even intimidation and threats of legal action resulting from the decision to fail students is also documented in the literature (Dudek *et al.*, 2005). Hunt (2014, p. 343) found that mentors feared making mistakes and “getting into trouble” if they failed to adhere to correct processes and procedures.

The feeling of guilt is also frequently experienced by mentors. Jervice and Tilki (2011) reported that mentors found it hard to fail students who were trying their best to achieve competence. Hawe (2003) stated that mentors found the experience of failing a student emotionally challenging since the student’s future career is jeopardised. Some mentors may feel that students’ failure reflects their own inadequacy to facilitate teaching and provide appropriate feedback (Dolan, 2003). They tend to blame themselves as being a bad mentor, perceiving it as a personal failure (Webb and Shakespeare, 2008). Failing a student may challenge their values of being caring (Luhanga *et al.*, 2008c). Black *et al.* (2014) found that the guilt mentors experienced included doubting their own competence as mentors

for not turning the failing student round, leading to physical manifestations of sleepless nights and exhaustion. Jervis and Tilki (2011, p. 584) added that the process of deciding to fail a student resulted in “considerable soul searching and stress”. Luhanga (2008c) identified that such feelings led to burnout, high levels of sickness and low retention rates.

As discussed earlier in section 2.2.1, assessing the practice component falls largely outside the control of the HEIs (Yorke, 2005) with mentors making unilateral decisions alone without the quality assurance measures employed by their academic colleagues. The literature provided evidence that mentors reported becoming isolated when their student demonstrated poor practice, exacerbated by loss of confidence and stress that made mentors reluctant to ask for help (Hunt, 2016). These decisions, Black (2014) added, are sometimes taken against the norms of a particular organisation or culture, and that it takes courage to stand up for what is right, which may exacerbate mentors’ fear or anxiety. The extent of making such difficult decisions were referred to by using powerful words such as an “act of bravery” (Luhanga *et al.*, 2008c, p. 8) or according to Hunt (2014, p. 102) requiring a “core of steel”, which reflected the enormity of the task that mentors face.

Mentors’ aptitude

Stakeholders in the scoping exercise pointed out issues around mentors’ capacity to perform the assessor role. Various factors influenced mentors’ abilities, and, as discussed earlier, lack of experience and inadequate preparation coupled with the pressure of service demands and low staffing levels played a significant part in mentors not feeling motivated to take up the role (Elkan and Robinson, 1995).

However, there are considerations as to whether mentors possess suitable personal characteristics to undertake the role since it was often externally imposed on them, as it is an expectation stated in the professional code of practice (NMC, 2008b, 2015, 2018d). This is intensified by the lack of choice or provision of protected time, as well as denial of remuneration or additional status (Nettleton and Bray, 2008). This, according to Baillie (1993), may result in some mentors having a negative

attitude towards students, jeopardising their ability to be effective mentors. In the new NMC standards (NMC, 2018a), there are no details provided to stipulate whether nurses will choose to become practice assessors or whether it will be imposed on them as is practiced currently, nor any indications whether the NMC expects time to be protected.

Robinson *et al.* (2012) argued that making a mentorship qualification an essential criterion for promotion in healthcare organisations coerced many nurses to undertake mentorship courses to enhance their job prospects, despite not having aptitude, desire or motivation to take up the role. This view was shared not only by mentors in the scoping exercise, but also by students who experienced mentors expressing these views to them. Robinson *et al.* (2012) proposed that the role should be developed into a specialist career pathway for those who want to specialise in nurse education. This echoed Nettleton and Bray's (2007) findings that making mentoring voluntary, where only mentors who have an interest in education became mentors, would improve practice. However, there will be a need for appropriate criteria and systems in place for mentor selection, and the ramification of this on having enough mentors for the number of students is likely to prove challenging.

The need to achieve performance targets focussing on a certain number of staff receiving mentorship qualifications may have also contributed to some unsuitable or newly qualified individuals becoming mentors. It is likely that such mentors would have minimal experience of assessment or understanding of the importance and implication of mentoring on public protection (Nettleton and Bray, 2008). In addition, Rutkowski (2007) pointed out, that when reviewing mentors' abilities to undertake the mentoring role, the literature does not consider the disparity between the preparation of mentors and that of lecturers as discussed earlier.

The stakeholders in the scoping exercise also identified the dual role to be a factor affecting a reliable and valid assessment of students. These views concur with published literature confirming that mentors' ability to comprehend the role is

further complicated by the dual role of being a mentor and assessor, which is exclusive to nursing in the UK and the ROI. The evidence suggested that mentors experience ambiguity about their role and responsibilities and that they understand their role to only be a supportive one (Wilson-Barnett *et al.*, 1995; O'Hara, 1996; Chow and Suen, 2001). In one study, Nettleton and Bray (2007) interviewed twenty mentors and none of them identified assessment as a central aspect of being a mentor. This perception did not change despite the NMC standards for mentorship preparation programme placing greater emphasis on the skills needed to assess competence, and the importance of ensuring that mentors are aware of their accountability for public protection through failing unsafe students (NMC, 2008a).

The confusion of the dual role is illustrated in the moral dilemma of assessing and judging a student at the same time as being their 'counsellor and friend'. This may create a conflict of interest, compromising mentors' objectivity when assessing performance (Neary, 1997) or when giving negative feedback (Bray and Nettleton, 2007). Moreover, stereotyping may further compromise the objectivity of assessment. Getting to know the students may lead to the 'halo' effect, which occurs when favourable knowledge of a student leads to making a more favourable judgement. Dennis (2007) explained that halo errors are persistent cognitive bias in subjective assessments of performance that occur when the assessor spends extended time with the student. This is of great importance in assessing nursing students in practice as Dennis (2007) warned that overstating students' performance mostly occurs when the assessor operates singly and knows that there is no cross-check on their judgement. Conversely, the opposite may occur which is termed the 'horn' effect, refers to the tendency to limit the overall assessment of an individual to a single negative attribute (MacDougall *et al.*, 2008).

Decision-making

Confidence to make decisions was also raised by many participants in the scoping exercise. Making a judgement to determine whether the practice observed fulfils the required standards is a challenging experience for mentors (Black *et al.*, 2014; Hunt, 2016). The professional accountability of formally assessing students in

practice entails responsibility for their actions and accepting the consequences of the decisions made (NMC, 2005; Rutkowski, 2007). Hence, the greater part of the decision-making process revolved around mentors judging if they could endure the anticipated challenges to their decision, rather than focussing only on whether the student performed to the required level (Hunt, 2016).

Many participants in the scoping exercise raised issues in relation to difficulties in assessing attitudes and behaviours. The NMC's (2010) holistic definition of competence (discussed in section 1.1.4) emphasises that possessing appropriate attitudes and behaviours are fundamental attributes for mentors to consider when assessing competence. Accordingly, mentors need to observe knowledge, skills and attitudes and make value judgements which can vary not only from person to person but from one situation to another (Cassidy, 2009). Therefore, it is important to note that uncertainty of decisions is influenced by the complexity of assessment in the inherently unpredictable and un-standardised real world (Govaerts and Van der Vleuten, 2013).

Nevertheless, mentors found making decision about the 'softer' aspects of competence such as empathy and compassion (known to be notoriously challenging to define or measure: Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016) to be difficult. This is different to judging competencies that focus on clinical skills such as administering an injection. Similarly, Fitzgerald *et al.* (2010) highlighted that the ability to give accurate feedback was a much deeper issue in relation to the professional values and behaviours, compared to the feedback on clinical skills. Hunt (2014) also stated that, to justify a fail decision, mentors tended to refer to unsafe practical skills and knowledge deficits because these were easier to substantiate.

There is also evidence in the literature that mentors rely on the subjectivity of their intuition, first impression and gut instinct when making decisions (Duffy, 2006; Black, 2011). Intuition is defined by Benner and Tanner (1987, p. 23) as "understanding without rationale". Tanner (2006) explained that intuition can be

characterised by recognising a pattern that makes a practitioner feel concerned about a situation in practice, therefore it has its credibility in clinical decision-making in nursing (King and MacLeod Clark, 2002; Smith *et al.*, 2004). Pattern recognition relating to student performance in practice can be useful, but this needs to be balanced with objectivity (Jervis and Tilki, 2011).

Objectivity of assessment decisions is informed by the official guideline from the HEIs (Payne *et al.*, 1997), but distancing the HEIs from clinical assessment makes it possible that the decisions made by mentors are informed by their own model of practice rather than the official guidelines. With the greater London area having a nursing workforce that has substantial numbers of overseas-trained nurses, there could be numerous models of practice being employed as well as cultural, ethical and moral differences. Allan (2010) found that mentors are ill-equipped by existing mentor preparation programmes to mentor overseas-trained nurses from culturally diverse backgrounds. This is even more relevant considering that some of the participants in the scoping exercise highlighted that their assessments are influenced by the way they were assessed themselves as students.

Making decisions about borderline students is particularly challenging and is manifested by mentors giving students the benefit of the doubt (Duffy, 2003; Gainsbury, 2010; Brown *et al.*, 2012). Robinson *et al.* (2012) found that the introduction of sign-off mentors may have contributed to more mentors allowing underperforming students to progress knowing that the final assessment would be made by a sign-off mentor experienced in assessment, resulting in a tendency for uncertain mentors to leave decisions about failure to the end of the course.

2.4 Identifying high impact barriers

The views shared by stakeholders in the scoping exercise echoed the wider national and international literature findings, demonstrating that the barriers to effective assessment in clinical practice are more than just a local issue and that research into practice-based assessment conducted in the local area is likely to be generalisable.

What emerges from this contextual review is that failing to fail and the factors contributing to it are significant issues in practice-based assessment, which calls for practical initiatives to deal with this complex problem. Exploring alternative ways that may make assessment in practice fit for purpose and evaluating its effectiveness is necessary. A pragmatic approach to identify areas that have highest impact on practice-based assessment was adopted; the underpinning premise is that identifying and modifying aspects that have big or high impact will result in significant improvements.

The term 'high impact' is widely used within the health services and associated with modernisation and service redesign. In nursing, the 'High Impact Actions' policy was launched in 2009 to improve areas where there is evidence of inefficiencies in care provision and poor patient experience. The intention is to focus on redesigning problematic areas that have the biggest impact on outcomes (Beasley, 2009).

Although they are not the product of academic theory, high impact actions are built on successes already achieved by frontline clinical teams (Fillingham, 2004). Their relevance to this research study is that they focus on significant gaps between current performance and best practice, with the aim to address the 'bottlenecks' by actively seeking them out, then redesigning the areas that cause them (NHS Modernisation Agency, 2004). Taking a similar pragmatic approach, a suggestion that can be put forward is to identify and act upon barriers that have a high impact on hindering practice-based assessment, and then evaluate if modifying these bottlenecks results in significant improvements.

When applying the principles suggested in the high impact actions process (NHS Modernisation Agency, 2004), Lack of time was the most frequently mentioned barrier to effective mentoring. Redesigning the assessment processes by, for example, evaluating whether providing protected time to alleviate this bottleneck would help mentors to function adequately in their role could be suggested. However, as discussed earlier, despite the DH (1999), the NMC (2008a, 2018e) and the RCN (2015) calls for protected time, enforcing such a change that includes

freeing mentors from their clinical duties, in already overstretched and understaffed clinical areas (NHS Employers, 2014), is considered an unrealistic and unfeasible research inquiry, as it is very unlikely to be endorsed by the health organisations due to the shortage in nursing manpower along with the associated financial implications of freeing mentors.

The second most frequently reported barrier in the scoping exercise was related to the ambiguity and unfamiliarity with the language used in the practice assessment documents, and the inability to determine what constitute competence. Therefore, the way competencies in the practice assessment documents are interpreted merits further analysis.

2.5 Chapter summary

This chapter has discussed the radical transformation that pre-registration nursing programmes went through in the UK that signalled a new and formal role for mentors in determining student nurses' fitness for practice. The discussion identified that inadequate preparation and confusion about the dual role resulted in concerns that mentors are not effective gate keepers. A scoping exercise conducted locally with stakeholders identified several factors affecting the quality of practice-based assessment. Ambiguous terminology and difficulties in identifying performance levels in practice assessment documents emerged as having a significant impact meriting further analysis. The next chapter presents an integrative literature review that systematically identifies and explores the literature in relation to understanding the terminology and levels of competency statements, and to determine how this influences fitness for practice decisions. The aim of the review was to systematically narrow down and identify the significant gap in knowledge, which this research study addresses.

Chapter 3: Integrative review

3.1 Introduction

The previous chapter provided a contextual review of the historic and current state of mentoring pre-registration nursing students in the UK and provided justification for further investigation to the issues related to the terminology of competencies and determination of appropriate levels. This chapter presents an integrative review (IR) that systematically summarises and synthesises empirical and theoretical literature and provides a broader and deeper insight into the quality of mentors' interpretation of competence in their assessment of nursing students. This IR has been published (Almalkawi *et al.*, 2018) but a fuller explanation of its design and results is presented here.

The IR presents the theoretical approaches underpinning this study: assessment *for* learning and authentic assessment. Key features of these concepts are discussed including how they influenced the study design. The chapter concludes with the study aims and research questions to address the gap in knowledge identified in this review.

3.2 Defining the term 'quality' in this review

This review aims to investigate the quality of mentors' interpretation of competency statements within practice assessment documents, thus it is important to define the term 'quality' in relation to this review. Dochy *et al.* (1990) described 'quality' as a value-laden term subjectively associated with what is good and worthwhile. Harvey and Green (1993) explained that like 'liberty' and 'equality', 'quality' is a slippery concept that is hard to articulate despite most people having an intuitive understanding of what it means.

The notion derives from the manufacturing industry that purportedly assesses a product against its stated purpose implying reliability and validity in the definitions, such as 'fitness for intended use' (Juran, 1999) or 'conformance with requirements' (Hoyer and Hoyer, 2001). Although this definition was intended for business to

evaluate how well a product performs for its intended use, it resonates in professional education where 'fitness for purpose' equates for quality with the fulfilment of a specific or stated outcome, and is usually based on the ability of an institution to accomplish its mission or for a programme of study to fulfil its aims (Harvey and Green, 1993).

On the other hand, the quality code set by the Quality Assurance Agency (QAA) emphasises two fundamental principles that should be evident in every assessment. First, assessment cannot be effective unless it is reliable (consistently performs its intended function) and valid (testing precisely what the examiners want to test). The second principle focuses on employing sound processes by ensuring that assessment is conducted with rigour, probity and fairness (QAA, 2012).

Stuart (2013) echoed the QAA principles stating that reliability, validity, feasibility and discriminating power are the four fundamental criteria of every effective assessment. Therefore, the concept of quality adopted in this IR (and the whole thesis) captures the extent to which mentors provide a 'fit for purpose' assessment (Juran, 1999), that employs reliable and valid strategies to distinguish and interpret different levels of practice and provides a reliable and valid judgement on whether a student is competent or not against set standards (QAA, 2012; Stuart, 2013).

3.3 The approach used to review the literature

Following from the previous chapter where barriers that have big impact on practice-based assessment were identified, the review of literature was conducted to gain deeper insight specific to the terminology and discrimination between performance levels in practice-based assessment criteria. From scoping the literature, it was evident that the concept of mentorship in nursing is very complex, reflected in the social, political, professional and conceptual frameworks that are intertwined in how student nurses should be assessed in practice. It was critical that the approach to reviewing the literature captured broad and 'authentic' data sources to understand such complexity, guiding the need to review experimental

and non-experimental research, policy, grey literature and theories in order to fully understand the phenomenon of concern.

Conducting a systematic review was considered for this study as this approach is considered at the top of the hierarchy for grading the quality of evidence (Oxford Centre for Evidence-Based Medicine, 2009). However, due to the heterogeneous and often small size studies relevant to the topic, it became clear that a systematic review, which according to Whitemore and Knafl, (2005) focuses on answering clinical questions and is exclusively conducted through a strict selection process and analysis of 'like' publications on the problem under study, was not an appropriate strategy. Whitemore and Knafl (2005) added that there is a misconception about systematic reviews when used to combine quantitative and qualitative studies, stressing that they should only be used for experimental studies.

The IR framework, on the other hand, is a midway between a literature review and a systematic review, capable of presenting varied perspectives of evidence (Russell, 2005). An IR does not employ summary statistics due to the heterogeneity of the studies but aims to summarise and compare data to allow for the achievement of general conclusions about the research problem. The strength of IRs is in their rigorous methodology and detailed search strategy in finding relevant evidence to answer the research question. Thus, they are considered to provide strong scientific evidence when strict methodological procedures are followed. Additionally, the distinctive feature of drawing conclusions from empirical studies and theories enhances the holistic understanding of the topic in question, creating a more well-rounded evidence review (Whitemore and Knafl, 2005).

According to Torraco (2005), IRs have the potential to produce a comprehensive interpretation of complex concepts and, relevant to the overall aim of this research study, "have direct applicability to practice and policy" (Whitemore and Knafl, 2005, p. 546). However, combining diverse data sources is complex and challenging, and there is little guidance available on how to conduct an IR. To address the challenges of combining diverse data sources, Whitemore and Knafl (2005)

modified Cooper's (1989) framework for systematic reviews and meta-analysis to make it suitable for IRs. Indeed, most of the guidance is predominantly provided by Whitemore and Knafl (2005), hence their framework is heavily cited in all IR publications as is the case in this review.

Accordingly, this IR follows a systematic and comprehensive approach for retrieval and synthesis of evidence employing Whitemore and Knafl's (2005) modified IR framework to facilitate the incorporation of both qualitative and quantitative evidence and provide a broad perspective linked to theoretical frameworks were appropriate. IRs focus on a process of five stages: 1) problem identification, 2) literature search, 3) data evaluation, 4) data analysis, and 5) presentation. Conducting these five stages properly, qualifies the results and allows for the identification of knowledge gaps regarding the phenomenon under study, and the identification of the need for future studies (Whitemore and Knafl, 2005).

3.3.1 Stage 1: Problem identification

The development of a well-specified review purpose and variables of interest intended for examination is the initial stage in IRs, which subsequently facilitates all other stages of the review (Whitemore and Knafl, 2005). Cooper (1998) noted that the problem identification process should include the development of conceptual and operational definitions of variables to be examined. Whitemore and Knafl (2005) explained that any IR can encompass an infinite number of variables, issues, or population, reiterating the importance of having clear research purpose.

The problems affecting the quality of mentors' assessment of nursing students in practice are complex and have been reviewed in chapter two. The intention throughout this study was to be pragmatic, aiming to make alterations and redesign 'bottlenecks' then evaluate the outcomes. The second most common issue identified in the scoping exercise after lack of time (see section 2.3) was related to the processes (summarised in Table 3.1 overleaf) as having a significant impact on the quality of the assessment process and outcome and this finding was mirrored in

the wider literature. Therefore, the process category formed the problem that merits further analysis in this IR.

Table 3.1: The processes category identified in the scoping exercise

Understanding/familiarity with the assessment documents
Understanding of competency/criteria
Understand what needs to be demonstrated to be worthy of a pass
Identifying levels of performance.
Accurate/constructive feedback
Inconsistency between assessors

The intended purpose was to explore the literature to identify the various factors influencing how mentors interpret different levels of competence and to evaluate the actual criteria or strategies mentors use to distinguish between different levels of competence and ensure their assessment is fit for purpose. Closely related, is the facilitation of constructive feedback.

3.3.2 Stage 2: Literature search

The second stage in the IR framework is the literature search. As with any type of review, a well-defined search strategy is central to enhance rigour in IRs, which necessitates a comprehensive search to identify the maximum number of eligible primary sources. This involves the challenging task of including all relevant literature related to the topic of interest (Cooper, 1982).

Additionally, to enhance the ability of readers to evaluate the search adequacy, explicit and systematic justification of the sampling decisions must be clearly documented in the method section, including search terms, the search strategies used, and the inclusion and exclusion criteria for determining relevant primary sources (Whittemore and Knafl, 2005). However, when defining the search question, despite Whittemore and Knafl (2005) emphasising the importance of having a well-defined literature search strategy, their framework did not discuss or provide clear guidelines regarding a search question format for IRs. This was evident in the variety of formats used in published IRs, ranging from those who used the Population, Intervention, Comparison and Outcome (PICO) format to others who use no identifiable format.

In the absence of a recommended format, and since the topic in this study has an educational rather than a clinical focus, the search question in this review employed the Best Evidence Medical and health professional Education (BEME) collaboration guidelines for reviews undertaken in medical and healthcare education. Parallel to PICO, which focuses on clinical questions (Polit and Beck, 2012), BEME focuses on health-related educational searching methods and recommends search questions where the queries can be broken down into Participants, Educational aspects and Outcomes (PEO). Defining the PEO elements used for framing the research question are presented in Table 3.2.

Table 3.2: The PEO model used for framing the research question	
PEO element	Defining the PEO element
Participants	Mentors and students, where mentors have responsibilities of assessing students' performance and determining fitness for practice.
Educational aspects	Practice-based assessment defined as the assessment which takes place in the practice settings in real life situations.
Outcome	Quality of the assessment. The term 'quality' (as defined in section 3.2) is used in this review to reflect reliable and valid strategies to interpret and distinguish different levels of practice and judge whether a student is competent or not against set standards.

The intention was to avoid having definitions that were too narrow at this stage, so the quality of the findings would not be impaired as this could be a threat to validity (Cooper, 1989). A broad identification of search terms was conducted through examining each essential subject component and using synonyms, alternative spellings, and related terms where applicable. For example, synonyms to the terms mentor and student were utilised to expand and include alternatives used by different healthcare professionals in different countries. Synonyms used for participants, educational aspects and outcome are detailed in Table 3.3 overleaf.

Table 3.3: PEO synonyms used to expand the essential terms			
Participants	Educational aspects	Outcome	Not
Mentor*	"Work based assessment"	Interpret*	Patient*
Assessor*	"Workplace assessment"	Language*	"Patient* assess*"
Preceptor*	"Workplace based assessment"	Terminolog*	"tool validat*"
Supervisor*	"Practice-based assessment"	"Level* of competenc*"	OSCE*
Trainer*	"Performance assessment"	"Level* of performance"	Simulate*
"Clinical educator*"	placement*	Discriminat*	Classroom*
Undergraduate*	"Clinical placement*"	Reliabl*, Valid*, Fair*,	
Student*	"Student placement*"	Robust*, Rigour*,	
Learner*	"Practice placement*"	Effectiv*, Accurat*,	
Trainee*	"education* measurement*"	Sensitiv*, Specific*	
Mentee*	"Practice document"	Capabil*, Competenc*	
Preceptee*	"Clinical competenc*"	Fitness*	
	"performance indicator*"	"Fitness for practice"	

Most of the databases were accessed through 'EBSCOhost' (Elton B. Stephens Co) as it has the advantage of facilitating searching several databases at the same time using one variation of wildcards and truncation symbols and automatic removal of duplicates. As shown in Table 3.3, the root of the keywords is used with truncation symbols (*) and quotation marks to search for an exact phrase. The wildcard symbol (?), used to replace a letter, was not required in this search. The combining commands based on Boolean operators was employed to join all alternative synonyms with OR. The term 'AND' was used to make the search more specific to PEO. 'NOT' was also used to exclude non-relevant terms such as patient assessment, classroom assessment, simulation or studies that focussed on tool validation.

Inclusion and exclusion criteria

Parameters, in the form of inclusion and exclusion criteria, were used to limit the search to only relevant results (Table 3.4 overleaf). The primary criterion for inclusion in the search was that articles must be related to any of the variables of interest presented in Table 3.3 that were identified in the processes category of the scoping exercise as discussed earlier (see Table 3.1). Articles were selected for review if they referred to mentors' interpretation of clinical competence or explored what strategies mentors employ to measure students' level of competence.

Table 3.4: Inclusion and exclusion criteria	
Inclusion	Exclusion
<ul style="list-style-type: none"> • Nursing. • 1986- present. • Practice/clinical based assessment. • Practice-based mentors undertook the assessment. • Tools and processes of conducting the assessment. • All publications explicitly related to the review questions and 'the processes' category in Table 3.1: (understanding/familiarity with the assessment documents, understanding of competency/criteria, understand what need to be demonstrated to be worthy of a pass, identifying levels of performance, accurate/constructive feedback, and inconsistency between assessors). 	<ul style="list-style-type: none"> • Pre- 1986. • Process of mentoring differs from the UK. • Faculty assessing practice, e.g. lecturers, clinical tutors, practice educators or clinical teachers. • Mentor role is exclusively advisory or facilitator. • Classroom/simulation/OSCE assessment. • Non-English. • Tool validation studies. • Personal opinion literature. • Other HCP students.

The search scope was to identify all articles published in English since 1986 (the formal introduction of mentoring in the UK) and related to mentors assessing students in practice settings in nursing. Literature from all countries was considered; however, articles were excluded if the practice-based assessment process differed from the nature of mentoring student nurses in the UK where mentors have the responsibility to determine students' fitness for practice. Therefore, studies were excluded if lecturers, clinical tutors, practice educators or clinical teachers carried out the assessment or the role of the practice mentor was that of an adviser or facilitator of learning only.

Articles were also excluded if the assessment was classroom-based or simulated practice such as OSCE (Objective Structured Clinical Examination). Since the NMC approved that up to 300 hours of practice can be simulated (NMC, 2010), such practice is likely be assessed by HEIs; furthermore, it would not reflect the authenticity of real-life practice assessment. Studies that introduced tools as a strategy to support assessment were included. However, tool validation studies were excluded, since their focus would be on reporting the reliability and validity properties of the tool itself rather than examining the quality of mentors' assessment.

Since an IR summarises past empirical or theoretical literature to provide a more comprehensive understanding of a specific phenomenon, or healthcare problem, the decision was made not to limit the search to peer-reviewed literature to allow retrieval of all related material including organisational and governmental publications and reports. However, personal opinion articles were considered not appropriate and were excluded.

Searching the databases

The literature search was conducted in September 2015 and followed an extensive and systematic approach across twelve electronic databases covering health and education related publications (Appendix 3). Medline, CINAHL Plus, PsycINFO, ERIC, ERC and AMED were searched combined through EBSCOhost, and the remaining databases (BNI, EMBASE, Cochrane, DARE, Joanna Briggs Institute and ETHOS) were searched individually.

Searching relevant education-health related journals (e.g. Nurse Education Today, Nurse Educator, Nurse Education in Practice) was conducted through their websites using subject-based search to identify literature not picked up by databases. Grey literature including professional body and DH databases were also searched. Additionally, the reference lists of the initially retrieved articles were searched manually as well as using facilities in Google Scholar and Science Direct to search for related papers. Authors of the retained studies were contacted by email to identify any relevant papers including unpublished literature.

Results

Based on the search terms and inclusion and exclusion criteria, the initial search resulted in 1910 hits retrieved from EBSCOhost and 451 hits retrieved from other databases and sources. Searching grey literature yielded 28 hits. After removal of duplicates, the total records identified were 1951. They were subsequently assessed for relevance based on title and abstract, resulting in 27 records retrieved for full-text review. Eight articles met the inclusion and exclusion criteria for final synthesis.

The process used to refine and evaluate each stage is presented in Figure 3.1 as PRISMA statement (Moher *et al.*, 2009).

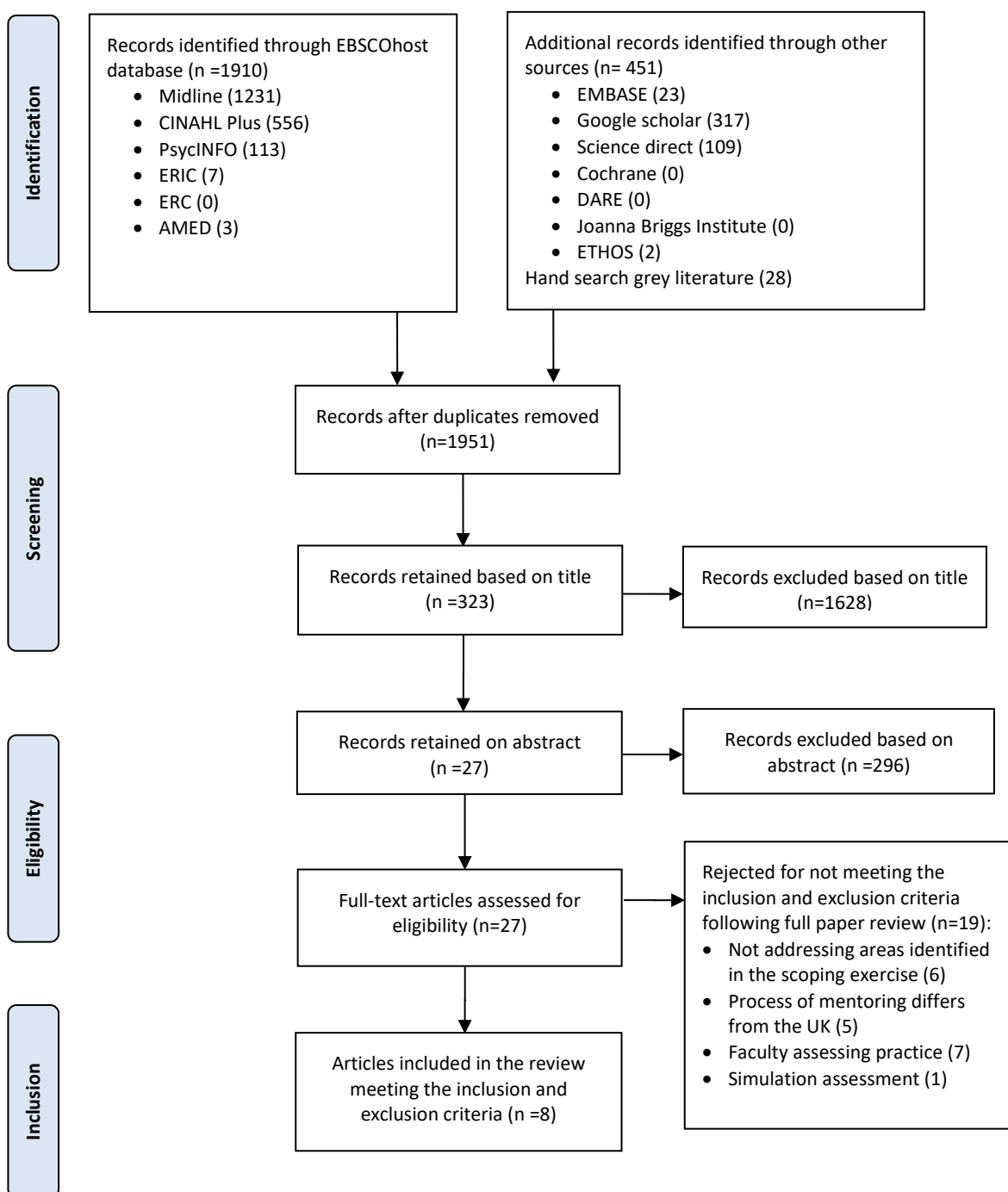


Figure 3.1: PRISMA flow chart of the literature search

3.3.3 Stage 3: Data evaluation

The third stage in the IR process is the data evaluation phase, with the aim to judge the quality of results and whether they are worthy of remaining in the data set (Cooper, 1998). The eight studies found to meet the inclusion and exclusion criteria were published between 2000 and 2012 and comprised one quantitative, one qualitative, five mixed methods and one literature review.

Four studies were conducted in the UK. The specific nature of mentoring in the UK, where mentors undertook the assessment in practice and made decisions about fitness for practice, is similar to the ROI only, from where the remaining four studies originated. Three papers from the ROI studies (Butler *et al.*, 2011; Fahy *et al.*, 2011; Cassidy *et al.*, 2012) were drawn from different phases of one mixed methods study and for the purpose of this review each study was reported separately since there were differences in the participants and methodology used in each study. The study by Neary (2001) summarised her PhD thesis (Neary, 1996), that was subsequently published in two parts in 2000 (Neary, 2000a; Neary 2000b). Since both publications referred to same study and reported the same outcomes, Neary (2001) was used in this IR.

Critical appraisal of the methodological features to evaluate the quality of studies in IRs is complex due to the inclusion of diverse primary sources. Whitemore and Knafl (2005) explained that no gold standard exists for evaluating and interpreting quality in IRs, and how quality is evaluated will vary depending on the sampling frame, for example, sources with similar design, calculating the quality score and incorporating them into the analysis may be optimal. However, if the primary sources include diverse sampling, Whitemore and Knafl (2005) suggest that it is reasonable to evaluate the methodological quality of the outlier primary sources to identify reasons for the discrepancy. Alternatively, similar to historical research, it may be appropriate to consider and discuss the quality of primary sources in the final report. Ideally, although complicated, consideration should be given whether to undertake a research design specific quality evaluation with an instrument for

each type of source and scores used as criteria for inclusion or exclusion (Whittemore and Knafl, 2005).

Since no gold standard exists, the decision was made to adopt the Mixed Methods Appraisal Tool (MMAT). The MMAT was designed to help overcome the challenges associated with appraising the methodological quality of studies with diverse designs (Pace *et al.*, 2012). It has the advantage of providing detailed and practical assessment of qualitative, quantitative and mixed methods research using one tool. The purpose of this tool is to allow for the concurrent appraisal of studies employing the most common methodologies and methods. To enhance consistency, guidelines for a few generic quality criteria are included in the tool (Pluye *et al.*, 2009; Pace *et al.*, 2012).

For each retained study, the methodological quality was assessed using the MMAT scoring metrics (Version 2011). The MMAT checklist includes 19 items corresponding to five methodological domains: qualitative studies, randomised controlled trials, non-randomised studies, quantitative descriptive studies, and mixed methods studies (Appendix 4). For quantitative and qualitative studies, the score can be the number of criteria met and scores vary from 25% (one criterion met) to 100% (all four criteria met). For mixed methods research studies, the premise is that the overall quality of a combination cannot exceed the quality of its weakest component. Thus, the overall quality score is the lowest score of the study components. The quality scores for the studies included in this review are presented within the data extraction sheet in Table 3.5.

The methodological quality of the qualitative studies was ranging from 50% to 75%. The most common criticism of qualitative studies was the researchers not addressing their influence on data collection. Similarly, the methodological quality of quantitative studies had the same range of 50% to 75%, influenced mainly by sampling and response rates. Since the studies achieved 50% and above, and in the absence of clear guidelines in the MMAT framework on the cut-off point for exclusion of studies, no study was excluded on the basis of its quality.

3.3.4 Stage 4: Data analysis

The goal of this stage of the IR is to interpret and synthesise the evidence from the primary sources. Data should be categorised and summarised into a unified and integrated conclusion about the research problem (Cooper, 1998). Whittmore and Knafl, (2005) reported that strategies for data analysis in IRs are the least developed areas, advising that methods developed for qualitative designs are particularly applicable to the IR method where the similar data are categorised and grouped together. The method consists of four steps: 1) data reduction, 2) data display, 3) data comparison and 4) data conclusion and verification.

Data reduction and display

The data reduction and display steps involve extracting and classifying the primary sources data included in the IR according to some logical system. The extracted data can then be converted into a display that assembles the data in the form of matrix or spreadsheet to enhance visualisation of patterns across all the primary sources and serves as a starting point for interpretation (Whittmore and Knafl, 2005). For the studies retained in this IR, data were extracted and summarised in a data extraction sheet (Table 3.5 overleaf). The data extracted included the conventional characteristics such as author, intervention type, study design, results, limitation and, as mentioned earlier, the MMAT quality score. Specific to this IR, theories cited in the studies were also extracted in order to provide an integrated analysis, which is one of the distinctive features of IRs.

Table 3.5: Data extraction sheet to describe the studies included in the integrative review								
Authors, year, and country	Intervention type	Study population	Study design	Outcome measure	Results/key findings	Limitations	Theoretical underpinning	MMAT score
Butler <i>et al.</i> (2011) ROI	Explore mentors' views and experiences of a competency assessment tool and process to assess nursing students.	Mentors (n=837) with overall response of 30% (n=255).	Quantitative: questionnaire survey.	Explore mentors' views and experiences of a competency assessment tool and process used to assess BSc student nurses' clinical competence.	<ul style="list-style-type: none"> 48% of mentors disagreed that the performance criteria were clear. 50% of mentors reported that the indicators did not provide a clear description of what is required in the competency assessment. Mentors experienced some difficulty identifying the required knowledge, skills and attitudes. Mentors found the language of the competency assessment tool difficult to understand. The competencies were regarded as broad, vague, open to interpretation and not sufficiently defined. 	Low response rate. Specific to one competency document. One geographical region in Ireland.	Competencies adopted Benner (1984) and Seinaker & Bell (1979).	75%
Cassidy <i>et al.</i> (2012) ROI	Evaluate mentors' views of assessing students using competency-based approach.	Mentors (n=16)	Qualitative: semi-structured interviews and guided focus group discussions.	Mentors' views and experiences of assessing competency of students.	<ul style="list-style-type: none"> Different interpretations of competency statements by mentors. Difficulty with the language used to describe performance criteria and the wording was not user friendly. Difficulty in assessing competence of soft skills (e.g. therapeutic relationships). 	Small sample size Specific to one competency document. One geographical region in Ireland.	Competencies adopted Benner (1984) and Seinaker & Bell (1979).	75%
Dolan (2003) UK	Investigate whether a revised system is an effective measure of clinical competence.	Nursing students, tutors and clinical mentors. Total not clear. (8 students submitted documents for analysis).	Mixed methods: qualitative (focus group) and quantitative (document analysis).	<p>Participants' views of the revised system.</p> <p>Compare documented evidence for consistency in the assessment process.</p>	<ul style="list-style-type: none"> Inconsistencies identified in interpreting competency statements (students, mentors and tutors), or what exactly is required. Inconsistencies in the amount of supporting evidence required by assessors despite guidelines. Written evidence is not a guarantee for competence, mentors signed it without reading it. Mentors needed more training in assessment process. 	No objective measure used as a comparison with the revised system. Content analysis findings not clear, mentioned other institutions' document as potentially useful without exploring what was useful.	None specified	50%

Cont. Table 3.5: Data extraction sheet to describe the studies included in the integrative review								
Authors, year, and country	Intervention type	Study population	Study design	Outcome measure	Results/key findings	Limitations	Theoretical underpinning	MMAT score
Fahy <i>et al.</i> (2011) ROI	Evaluate students' and mentors' experience of clinical competence assessment.	Focus group: 13 Students/ 16 Mentors. Survey: 232 Students/ 837 Mentors	Mixed methods: qualitative (focus group) and quantitative (Survey).	Students' and mentors' experience and views on the assessment process, structure and content of the assessment document.	<ul style="list-style-type: none"> Students and mentors reported difficulties with the language used in the document. The language lacked clarity and required defining: too broad, vague and open to interpretations. Mentors understood the language better than the students but still had difficulty in making sense of the competence requirements. Students felt challenged to figure out what is expected of them and preferred more specific competencies. 	Low mentor response rate. Specific to one competency document. One geographical region in Ireland. Further clarification about recruitment of mentors needed.	None specified	75%
Girot (2000) UK	Examine if there is a difference between diploma and degree level competence and if it is possible to measure it.	NA	Reflective literature review.	Whether there is a difference between levels of competence and whether is it possible to measure it.	<ul style="list-style-type: none"> Clarification of academic achievements in practice assessment is needed. There is a paucity of research in relation to different abilities and expectations in practice. Assessment of practice is fraught with difficulties in defining and measuring expectations in the real word. Assessment strategies not effective in identifying poor performance, vague terms and documents lacked clarity. Need to enhance practice assessors' verification of achievements in practice. Few assessment tools allow the diversity and constraints of practice environment to be articulated. Difficulty articulating high levels of achievement. 	Not an empirical study.	Bloom' (1956) taxonomy for the cognitive domain.	50%
Heaslip and Scammell (2012) UK	Explore if using a grading tool improves reliability of mentors' judgements of students' levels.	Convenience sample of students (n=107) and mentors (n=112).	Mixed methods: questionnaire survey, fixed and free response questions.	Explore if using a grading tool improves reliability of mentors' judgements of students' levels.	<ul style="list-style-type: none"> The grading tool enabled mentors to become more discriminating in allocation of grades. Mentors were confident in grading practice but not confident in awarding a fail grade. Students reported inconsistency of mentors' use of the descriptors. Inconsistency in mentors' and students' perception about the amount of feedback provided 	Convenience sample from one institution, difficult to generalise.	Grading of competencies adopted Bondy (1983)	50%

Cont. Table 3.5: Data extraction sheet to describe the studies included in the integrative review								
Authors, year, and country	Intervention type	Study population	Study design	Outcome measure	Results/key findings	Limitations	Theoretical underpinning	MMAT score
McCarthy and Murphy (2008) ROI	Explore the use of reflection to clinically assess students.	Mentors (n=970) with overall 49% response rate (n=470)	Mixed methods: questionnaire	Reflection on practice. Framework to distinguish levels.	<ul style="list-style-type: none"> Inconsistencies with mentors' understanding of the assessment strategies. Assessors have their own interpretation of competency leading to various approaches used by assessors to assess students in clinical practice. 	One university. Mentors' responsibility to assess is recent. Qualitative responses may have added value.	Reflection models (Gibbs 1988 and Schon 1983). Bloom' (1956) taxonomy	50%
Near (2001) UK	Clarify students' and mentors' understanding of their role in the assessment process and their perceived preparation for it.	Interviews: students (n=70), mentors (n=80). Questionnaire student (n=155).	Mixed methods: semi-structured interviews and questionnaire	Clarify students and mentors understanding of their role in the assessment process and their perceived preparation for it.	<ul style="list-style-type: none"> Students experienced variations in practice placements concerning both what assessment criteria should be and what should be assessed. Students and mentors rarely knew the detail of the assessment criteria, or how to interpret them. Students and mentors negotiated their own objectives to cope with the messy language and 'ticking the boxes' to 'keep the college happy'. Students appreciated constructive comments in support of grades and considered that mentors needed to devote more time to the actual assessment and feedback process. Inconsistencies in how grades were awarded, for example, never giving an 'A' on principle or the use of arbitrary criteria. The college grading systems varied, and the meaning or value of the particular grading scheme was not clear. 	Methodology and analysis not clearly stated.	Taxonomies (Benner 1984, Dreyfus and Dreyfus 1980, Seinaker and Bell 1979 and Stake 1977) Stake's (1983) responsive evaluation model	50%

Data comparison

The next step in data analysis is data comparison which involves an iterative process of identifying and grouping similar variables to identify themes and relationships and provide clarity to the empirical and theoretical support. Accordingly, the primary sources in this IR were examined carefully to identify emerging themes with the same meaning. This resulted in identifying three themes:

- i. Difficulties with interpreting the language used within competencies
- ii. Difficulties distinguishing between different levels of competence
- iii. Difficulties articulating feedback regarding developmental needs to students

i. Difficulties with interpreting the language used within competencies

In the category related to difficulties in the language used to describe competencies in the assessment document, six studies (Neary, 2001; Dolan, 2003; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Cassidy *et al.*, 2012) reported this problem.

Three of the papers originated from the ROI and reported on separate parts of a multi-part study conducted by a team that included Cassidy, Butler and Fay. Two of the papers reported on one phase each of a two-phase study incorporating a qualitative (Cassidy *et al.*, 2012) and a quantitative (Butler *et al.*, 2011) phase, which explored mentors' views and experiences of the competency-based approach. The third paper by Fahy *et al.* (2011) reported the final mixed methods study to explore the views of both mentors and students.

In their first phase, Cassidy *et al.* (2012) used focus groups with 16 mentors to explore their perspectives on the content of a clinical competency assessment process used to assess BSc nursing students. In this study, mentors reported that the wording of the document was not user-friendly, and they had difficulty in describing performance criteria. This caused different interpretations of competency statements. The second phase of the study was reported by Butler *et al.* (2011) and involved surveying all mentors (n=837) in one region to explore their

views and experiences of a competency assessment tool and process used to assess BSc student nurses' clinical competence. Of the 255 mentors who responded (30% response rate), 48% of mentors did not agree that the language used in the competency assessment tool was easy to understand and 17% were undecided. They felt challenged to translate competency statements into assessable criteria, regarding them as broad, vague, open to interpretations and not sufficiently defined.

The report from Fahy *et al.* (2011) explained how they used focus groups with 16 mentors and 13 students in the first phase of the final mixed methods study to ascertain their views on the assessment process and documentation as well as informing the development of a questionnaire. The developed questionnaire was then used in the second phase of their study to explore the views of students (n=232) and mentors (n=837) on the structure and content of the competency assessment tool⁵. They were subject to descriptive analysis using SPSS with findings demonstrating difficulty in the language used, describing it as broad, vague and open to misinterpretation. Both students and mentors agreed that the knowledge, skills and attitudes required to complete each competency lacked definition and clarity.

Similar findings were reported in a study in the UK by Dolan (2003) who used focus group interviews with students, lecturers and mentors to examine how mentors use assessment strategies, and whether a devised system is an effective measure of clinical competence. All the groups reported inconsistency in the way competency statements are interpreted. Lecturers in this study acknowledged these differences and some felt that although assessment is based on set standards, "students need to accept that mentors may have their own interpretations of competence" (Dolan, 2003, p. 136).

⁵ Although the number of mentors participating matched those used in Butler *et al.*, (2011) and Cassidy *et al.*, (2012) studies, the author did not report that they were the same sample.

McCarthy and Murphy (2008) conducted another study that explored the extent to which mentors use devised assessment strategies to assess students in the ROI. A 24-item questionnaire using a Likert scale with two additional open-ended questions was administered to 970 mentors from one university and achieved a 48.5% response rate (n=470). Using descriptive statistical analysis, mentors reported that they had their own interpretation of competencies rather than relying on the pre-determined programme competencies, leading to various approaches used by mentors to assess students in clinical practice.

Further evidence was provided by Neary's (2001) mixed methods study, which was based on a PhD thesis she completed in 1996 examining students' and mentors' thoughts about the nature of continuous assessment of competence. Her objective was to clarify how mentors understand their role in the assessment process and their perceived preparation for it, with a final intention to develop an assessment tool. Semi-structured interviews with 70 students and 80 mentors were conducted to gather data not only used to explore participants' thoughts and beliefs, but also to design the main questionnaire that was used to collect data from 300 students and 155 mentors at various stages of a three-year course. Data indicated that students and mentors rarely knew the detail of the assessment criteria, or how to interpret them claiming that they were 'too objective bound' to suit the college academic requirements, using words to describe the assessment document such as 'very confusing', 'vague', 'ambiguous' and 'jargon'.

Documentation was also considered problematic with significant time and effort spent on making sense of the competencies in order to complete the assessment documentation, rather than to assess the competencies themselves (Neary, 1996). Hence, both mentors and students often found themselves trying to fit the objectives to the students' performance. As a strategy, they negotiated their own objectives to cope with the 'messy' problems of the practice assessment document, while "ticking the boxes to keep the college happy" (Neary, 2001, p. 5).

ii. Difficulties distinguishing between different levels of competence

For the category related to the problems associated with mentors' and students' ability to distinguish between different levels of competence, six studies (Giot, 2000; Neary, 2001; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Heaslip and Scammell, 2012) reported this problem.

In Butler's *et al.*'s (2011) questionnaire, mentors (n=837) were asked to rate their level of agreement on whether the assessment tool provided clear description of what was required; 50% reported that the indicators did not provide a clear description of what is required in the competency assessment. Mentors found that clinical skills were easy to identify, but they experienced difficulty in identifying the required attitudes, and only 36% agreed that the competency assessment framework adequately assessed students' clinical competence. Fahy *et al.* (2011) added that although mentors had difficulty in making sense of the competence requirements, students felt challenged to figure out what was expected of them and that they preferred competencies to be more specific.

The literature review conducted by Giot (2000) revisited the debate on the meaning of 'competence', particularly looking at the expected level of competence at both diploma and degree levels. The review examined whether there is a difference in their level of competence and if it is possible to measure it. The author explained that what distinguished diploma from graduate level is restricted to the cognitive domain and the distinction of practice outcomes for different levels is less clear as the expected outcomes are the same for both levels, arguing that the notion of having one level of competent practitioner with two different academic awards can no longer be justified. Giot acknowledged the paucity of, and need for, empirical research to compare different outcomes in practice.

Giot (2000) also reported that assessment of practice has been fraught with difficulties, not least the attempt to define expectations and ultimately the measurement of these expectations in the real world of practice, concluding that problems exist with the strategies, the tools and the mentors using them,

reinforcing that much more work needs to be done to produce a tool that presents a reliable and valid way of determining the achievement of fitness for practice. Although the author did not provide the methodology for conducting the literature review which may influence its rigour, the review provided relevant reports that contributed to the debate.

Heaslip and Scammell (2012) provided similar themes in a study that evaluated a tool developed to grade practice that was based on the work of Bondy (1983). A questionnaire was administered to 107 students and 112 mentors in one institution to investigate their experiences using a grading tool. The questionnaire collected data via quantitative (yes/no and 5-point Likert scales), qualitative (open questions) and multiple responses questions (ticking all that apply). Chi-square analysis using SPSS was used to explore the significance of results. The qualitative data were subjected to content analysis using conceptual analysis methods.

Heaslip and Scammell's results indicated that most mentors were confident in grading practice and found the tool useful in facilitating assessment of students' practice. This was reinforced by the qualitative data, which showed that mentors liked being able to allocate a grade to student performance as the level descriptors provided a benchmark from which students could monitor their progress. However, mentors identified some confusion regarding the use of the tool, which was reinforced by the students indicating inconsistency in mentors' use of the descriptors.

What was striking is that, despite mentors being confident to grade practice using the assessment tool, some mentors indicated that they were not confident in awarding a fail grade. This might relate to the reasons explored in the previous chapter more than the quality of the tool. Heaslip and Scammell (2012) concluded that their study appeared to show that assessment tools with more discriminatory grading systems and clear descriptors are helpful and welcomed by mentors, but this did not diminish the responsibility to identify failing students and act appropriately. The authors supported the recommendation outlined by Gray and

Donaldson (2009) for further research into the development of and testing of grading criteria or rubrics.

McCarthy and Murphy (2008) also reported that it was difficult to determine how mentors judged students' progression or level of performance. Of the 470 mentors who responded to the questionnaire, less than 50% referred to the tool introduced to assist students and mentors to distinguish levels of learning during the student assessment process, and 54% of mentors based their judgement decisions on assessing practical skills rather than competence. Mentors also encountered difficulties with guiding students through reflection but agreed that reflection is a helpful strategy for students to self-assess their own practice.

Similarly, Neary (2001, p. 6) reported that it was not clear that students and mentors were fully aware of the meaning or value of the grading schemes used and identified inconsistencies in how grades were awarded, for example, "never giving an 'A' on principle" or "the use of arbitrary criteria" and "inappropriate personal opinions" were mentioned during group interviews. Hence, the students experienced variations in practice placements concerning both what the assessment criteria should be and what should be assessed. Additionally, even though students were keen to use criteria-referenced assessment, they criticised the assessment booklets as being "burdensome" with insufficient scope to reflect student performance accurately and helpfully, describing the pre-set objectives as inflexible and seldom reflected real-life situations (Neary, 2001, p. 6).

iii. Difficulties articulating feedback regarding developmental needs to students

This category related to the provision of constructive feedback to identify developmental needs and show students how to improve their weaknesses or build upon what they do best. Two studies (Neary, 2001; Heaslip and Scammell, 2012) reported this problem.

Heaslip and Scammell (2012) stated that development of competence depends upon students receiving formal feedback. Thus, feedback is part of the assessment

process that supports assessment decisions as well as facilitating learning by enabling students to identify their strengths and weaknesses. Their study identified a lack of transparency in the feedback provided to students. Of the 112 mentors and 107 students who responded to the questionnaire, the majority of mentors (92%) believed that they provided feedback throughout the placement. This is in contrast to 57% of students stating that they only received feedback at the end of placement and 13% of students indicating that they did not receive feedback at all. Differing perceptions between mentors and students were also evident in their views whether the feedback provided reflected the grade awarded. Most mentors (89%) indicated that their feedback matched the grade awarded, whereas only 61% of the students perceived this to be the case.

Further examination of feedback was also reported by Neary (2001) in that students appreciated immediate feedback to enhance the formative assessment, but also valued constructive comments in support of grades. In this context, what was important, according to Neary (2001), is not only the frequency of the feedback but its nature and quality as well, which seemed to vary among mentors; consequently, students considered that assessors needed to devote more time to the actual assessment and feedback process.

Theoretical underpinning reported in the studies

Only Neary (2001) adopted the concept of 'responsive assessment' as a theoretical framework for her PhD. The remaining studies incorporated in this IR loosely referred to taxonomies used in assessment tools, mainly Benner's (1984) framework, and the use of reflective practice in the context of assessment.

Taxonomies

Several taxonomies were used in the literature to determine the different levels of practice. In the studies incorporated in this IR, reference to both Benner's (1984) 'novice to expert' stages and Seinaker and Bell's (1979) experimental taxonomy was made in four studies (Neary, 2001; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Cassidy *et al.*, 2012). Bloom's (1956) taxonomy was cited in two studies (Giot,

2000; McCarthy and Murphy, 2008), while Heaslip and Scammell (2012) referred to Bondy (1983). Bloom's taxonomy was reviewed in relation to the cognitive domain and focussed on learning only, therefore it will not form part of this review.

Benner's (1984) model of skills acquisition, commonly termed 'novice to expert', was based on a theory developed by Dreyfus and Dreyfus (1980) that defined five progressive stages that nurses may pass through on the way to achieving expert status namely: novice, advanced beginner, competent, proficient and expert. According to Dawson (2006) the model was initially adopted in nursing in the UK and referred to the learning outcomes specified in regulations by the UKCC (1989), resulting in most HEIs using clinical competence assessment tools based on Benner's (1984) framework. However, Calman *et al.* (2002) reported in their study that by 1997 most programmes had rejected Benner's approach describing it as unworkable, based on feedback from practice assessors and students who had difficulty in understanding the wording or applying the items.

A plausible explanation as to why Benner's (1984) model was not fit for purpose could be that Benner's research used qualified nurses, hence, her correlating of the novice stage to the initial qualification of a nurse rather than undergraduates was overlooked. Daley (1999) acknowledged that Benner's taxonomy has implications for post-registration education, but the UK policy on nurse education failed to identify that when the framework was adopted for pre-registration programmes. Benner viewed the newly qualified nurse as an "advanced beginner" which differs from how her framework ended up being employed to underpin pre-registration education. Second, with reference to the difficulty in defining the term 'competence' (discussed earlier in section 1.1.4), the competent stage is placed in the middle of Benner's framework, and at that level practitioners are expected to function safely, clearly implying that to be competent, one is not required to be proficient or expert but rather to be 'good enough' (Eraut, 1994), which is problematic and too ambiguous to measure (Yorke, 2005).

In the ROI, the studies appear to report that the Irish National Board competency statements are still based on Benner's (1984) framework. Steinaker and Bell's (1979) experiential taxonomy was also used to reflect the level expected in a process similar to the UK. The process includes a preliminary interview to agree the criteria for assessment, an intermediate interview to appraise progress and a final interview, where a mentor reviews the achievement of the agreed assessment criteria to pass or fail the student. In addition, Steinaker and Bell's taxonomy is allegedly used to guide students to use reflection on action as a clinical assessment strategy through using Gibbs's (1988) or Schon's (1983) reflective cycles. Nonetheless, the evidence has already been presented in this IR and remains consistent in confirming that mentors found defining the knowledge, skills and attitudes challenging (Butler *et al.*, 2011; Cassidy *et al.*, 2012).

McCarthy and Murphy (2008) indicated that less than 50% of mentors have clear knowledge and understanding of Steinaker and Bell's taxonomy or have used it to help them during the assessment process. With such a significant percentage of mentors declaring they were not using the framework, it is difficult to determine how mentors judge progression or level of competence. Moreover, McCarthy and Murphy (2008) reported that the majority of mentors acknowledged that reflection is a helpful strategy for students to self-assess their own practice; however, they encountered difficulties in guiding students through reflective cycles. More importantly, large numbers of mentors reported that they did not have the knowledge or skills to undertake reflection-on-action as an assessment strategy, and that reflective notes did not influence them when they assessed clinical performance. Maybe this is to be expected since reflective cycles play an important part in facilitating learning and do not explore the context of assessment, which may create tension.

On the other hand, Bondy's (1983) framework described a skills escalator structure that identified five levels of performance a student should be achieving in practice at given points in the programme. In this escalator, students can move up and down levels of achievement depending on the type of placement (Aston *et al.*, 2010).

Bondy's model was used to assess pre and post registration competence, hence has been criticised by Buckingham (2000) for being vague when assessing specific tasks. Indeed, in a follow up study, Bondy (1984) investigated the effect of the criteria on the accuracy and reliability of assessing students' clinical performance. In analysing the grades awarded to students, she found that the lowest grades were awarded for the psychomotor behaviours and the highest were awarded to the affective type behaviour. Bondy (1984) suggested that it appeared that attitudes and behaviour were assessed more leniently as they were perceived as more abstract and open to interpretation.

For her PhD thesis, Neary (1996) attempted to expand on the debate that while mentors recognised the importance of assessing students' competence in practice, trying to categorise learning objectives into various achievement levels was problematic. Her argument was that innovative assessment strategies are needed to assess 'real life' situations with intended and unintended learning outcomes. She used a conceptual model called 'responsive assessment' that shifts the focus away from routine application of standardised assessment measures and toward understanding the immediate educational needs of students. The responsive evaluation model measures the effectiveness of educational programs in a situational context by observing students' responsiveness to main issues and problems.

Responsive assessment, according to Neary (2001), accommodates for the fluid nature of clinical settings and therefore, in placements of varying complexity, it is reasonable to expect fluctuating levels of student performance. Responsive assessment assesses students' competence in a situational context, which fits better with developmental planning to evaluate personal growth rather than traditional assessment processes. The mentor role changes by requiring two actions: description and judgement. Thus, assessments become a process for both describing the students and judging the merit and worth of their performance. In this model, objective criteria for performance would not be the sole measure for comprehensive practical assessment.

What is problematic in Neary's responsive assessment model is that allowing mentors the right to judge according to the situation undermines the importance of judging against objective criteria, considered to be the most important pillar and core of defining reliable and valid assessments. Equally, it is crucial to recognise the influence of mentors' differences in their professional experience and personal beliefs and values that shape their responses in dynamic clinical environments. Thus, competency cannot be described with sufficient precision to guarantee equivalence between individuals and across situations. In addition, the responsive assessment model was partially based on Stake's (1977) model which was criticised in earlier literature as being complex to operationalise and too difficult to comprehend (White, 1990).

Conclusion drawing and verification

This is the final step of data analysis in the IR, which moves from describing patterns and relationships to higher levels of abstraction and synthesis into an integrated summary of the topic or phenomenon. This IR examined the literature identified through a rigorous search strategy as relevant to the intended purpose of exploring clarity of competency statements and levels.

Methodologically, research into assessment of competencies needs to be stronger. Evaluation of the studies using the MMAT (Pace *et al.*, 2012) indicated the research quality of the studies included in the review was moderate. In general, there was a lack of a clear philosophical perspective underpinning the research designs. Theoretical underpinning was limited to making reference to taxonomies and reflection (presented in Table 3.5). Only Neary (2001) discussed other theories including the theory of 'responsive assessment'. Policy drivers (discussed in detail in the previous chapter) that had a significant influence on shaping practice-based assessment of pre-registration nursing students were evident in all the papers.

The descriptions of 'competence' need improving. The majority of studies reported that understanding the language used to describe competencies is challenging (Neary, 2001; Dolan, 2003; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et*

al., 2011; Cassidy *et al.*, 2012). These findings were also supported in other studies (Brown, 2000; Duffy and Watson, 2001; Norman *et al.*, 2002; Duffy, 2003; Scholes *et al.*, 2004; Miller, 2010). Difficulty in translating and applying assessment outcomes into observable practice activities may result in mentors and students having no confidence in practice documents (Norman *et al.*, 2002; Scholes *et al.*, 2004), with both mentors and students spending significant time trying to work out what the competency statements mean rather than assessing the student against them (Neary, 2000; Scholes *et al.*, 2004). Subsequently, mentors develop their own interpretation of competence leading to various approaches used to assess students.

The studies reviewed also reported problems in discriminating between different levels of practice and identifying the benchmark of what constitutes a pass or a fail (Giot, 2000; Neary, 2001; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Heaslip and Scammell, 2012). Consequently, this complicates mentors' judgement of performance (Norman *et al.*, 2002; Cowen *et al.*, 2005). This, in part, is directly related to the complexity and lack of consensus on what 'competent' really means, but there is also evidence that mentors have differing views about what is considered as an acceptable standard of competence a student needs to meet in order to pass (Cassidy, 2009).

Mentors had difficulty in making sense of the competence requirement, which may have led them to interpret the concept of competence differently; one mentor may pass a student while another may fail the same student when assessing the same competency, compromising interrater reliability, defined as the degree of consistency between two or more raters (Gray *et al.*, 2017). Thus, students experienced variations concerning what assessment criteria should be and what should be assessed (Neary, 2001). They felt challenged to figure out what was expected of them and favoured more specific competencies (Fahy *et al.*, 2011).

Assessment of competencies can vary from mentor to mentor and HEI to HEI. As a coping mechanism, mentors tended to negotiate their own objectives and learning

outcomes to deal with the burden of the academic jargon and filling in the forms just to satisfy the HEI (Neary, 2001; Bray and Nettleton, 2007; McCarthy and Murphy, 2008). Hence, when it came to justify their decisions, mentors struggled to prove their concerns were valid (Duffy, 2003; Gainsbury, 2010; Brown *et al.*, 2012). The problems may have risen from the fact that, despite all programmes stemming from the same regulatory body, HEIs are given the freedom to interpret, design and produce their own documents to meet their own needs, and these contained educational terminology unfamiliar to clinical practitioners (Rutkowski, 2007; Helminen *et al.*, 2016).

Neary (2001, p. 4) found that mentors often try to fit the objectives to the student's performance, claiming that assessment strategies were "too objectively bound" to suit the HEI academic requirements rather than providing sufficient scope to reflect students' performance accurately. This may be as a result of the level of abstraction used to construct the proficiency statements, so they accommodate a diverse range of situations (Scholes *et al.*, 2004). In an earlier study, Gerrish *et al.* (1997) analysed documents from nursing and midwifery programmes at different academic levels and found that assessment of practice varied across universities. In particular, different methods and criteria were used to interpret professional and educational frameworks (Gerrish *et al.*, 1997). Therefore, it came as no surprise that Dolan (2003) reported inconsistencies in the way students, mentors and university lecturers interpret competency statements, both within and between groups.

Students need to feel that mentors have assessed them fairly and competently. The notion that students, and possibly mentors, find reassurance if specific competencies to be assessed are identified is supported in an earlier study where students believed that more specific competencies would improve assessment tools (Calman *et al.*, 2002). This may also explain why large numbers of mentors focussed predominantly on assessing practical skills as a substitute for competence (Neary, 2001; Dolan, 2003; McCarthy and Murphy, 2008; Hunt, 2012). McCarthy and Murphy (2008) added that the absence of recognisable standards that indicate students' levels of performance could lead to students being educated to meet

minimum competency levels or according to Neary (2001) never given the maximum grade on principle or mentors using arbitrary criteria and inappropriate personal opinions.

Insufficient time for assessment affects the quality of the assessment decision. Certainly, insufficient time to dedicate to the assessment process (Dolan, 2003), coupled with the complexity of defining the term 'competence' (Cowan *et al.*, 2005), and the dynamic nature of nursing practice (Neary, 2001; Cassidy, 2009) contributes to this problem. However, the empirical evidence supports that the language of the competency statements is vague, open to interpretation and difficult to translate into assessable criteria is undisputed, stressing the need for clear and unambiguous language of competency assessment documents (Butler *et al.*, 2011; Fahy *et al.*, 2011). Others argue that a national collaborative approach to develop a uniform competency assessment document may enhance understanding (Norman *et al.*, 2002; O'Connor *et al.*, 2009; Cassidy *et al.*, 2012). Fahy *et al.* (2011) supported such collaboration with specific emphasis on improving the language and terminology used.

On a regional scale, a collaborative approach developed a unified practice assessment document that had been in operation in London since 2014. The Pan-London Practice Assessment Document (PLPAD) created consistency with use of a single document for use around London and has been well received by mentors, students, and academic staff (Baillie *et al.*, 2016). However, The PLPAD adopted an achieved/not achieved classification, therefore interpreting and differentiating different levels of performance through deciding what constitutes competent continues to be subjectively decided by mentors. Additionally, the achieved/not achieved classification does not give credit to those exceeding 'achieved'.

This IR confirms that practice assessment tools need improving. As discussed earlier the academic component of nurse education seemed well-versed in distinguishing between levels and developed criteria of expectations at different academic levels, and assessment tools need to be specific and clearly articulated to guide both

academics and students. In clinical practice, however, despite the likelihood that mentors have less experience in assessment than their academic colleagues, there is a paucity of tools with meaningful criteria articulating the distinction between expected outcomes (Donaldson and Gray, 2012).

Tools that allow for the diversity and constraints of practice-based assessment to be articulated in relation to specific level descriptors may support mentors in determining students' level of achievement, and in providing constructive feedback that matches the performance. This would also support students by emphasising the value of self-assessment (Giot, 2000). This argument was supported by Fitzgerald *et al.* (2011) who reiterated that the lack of clear criteria against which students' performance can be judged, not only influences the accuracy of completing students' documents, but also reduces the likelihood of mentors offering constructive feedback that truthfully reflects the level of practice to move students forward.

Neary (2001) evaluated the various grading tools ranging from 'A–D', 'satisfactory–poor', 'pass/fail' and 'achieved/not achieved' and found that it was not clear that all who were involved were fully aware of the meaning or value of the particular grading scheme used. This is echoed in a study by Norman *et al.* (2002) who raised doubts about the reliability, validity and sensitivity of the available assessment tools in their ability to discriminate between competent and incompetent students.

Heaslip and Scammell's study (2012) confirmed these findings and reported that assessment tools that use more discriminating grading systems (as opposed to pass/fail) were helpful and welcomed by mentors. Nonetheless, although grading tools allowed mentors to become more discriminate in grading practice and feedback provision, the decision to award a failing grade remained problematic, another reminder of the complexity of mentors' confidence in making the decision to fail students.

Norman *et al.* (2002) proposed a need for a national assessment tool to facilitate the development of robust training for mentors that would transfer across clinical placements. Both Fahy *et al.* (2011) and Girot (2000) echoed this proposal and recommended that HEIs and health service partners should share the responsibility in reviewing and developing competence assessment strategies, and that any new model or approach to competence assessment must incorporate the views of both students and mentors.

3.3.5 Stage 5: Presentation

The final stage in Whitemore and Knaf's (2005) IR framework is to summarise the evidence from the primary sources to support conclusions that demonstrate a logical chain of evidence. Ideally, the results capture the depth and breadth of the topic and contribute to a new understanding of the phenomenon of concern and implications for research.

The current understanding of mentors' assessment of students suggests that there is no single uniform method of assessing clinical competence in pre-registration nursing education. This leads to inconsistency and difficulty in interpreting the language used to describe competencies in the practice assessment document (Neary, 2001; Dolan, 2003; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Cassidy *et al.*, 2012).

Examining the factors contributing to this problem shows that the term 'competence' is poorly defined. This is further complicated by too much freedom being given to HEIs to design their own documents which contain educational terminology unfamiliar to clinical practitioners. The need for a collaborative approach to use clear and unambiguous language for competency assessment was recommended by several studies (Neary, 2001; Norman *et al.*, 2002; McCarthy and Murphy, 2008; Fahy *et al.*, 2011; Heaslip and Scammell, 2012).

Difficulties in discriminating between different levels of practice is also demonstrated in this review (Girot, 2000; Neary, 2001; McCarthy and Murphy,

2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Heaslip and Scammell, 2012). The inability to discriminate compromises mentors' judgement of performance and reduces the likelihood of mentors offering constructive feedback that truthfully reflects the level of practice to move students forward.

Examining the literature identified that tools with articulated competency statements in relation to specific level descriptors may support mentors in determining levels of achievement and provide feedback that matches the performance, as well as facilitating students' self-assessment. This finding takes into consideration Fahy *et al.*'s (2011) and Girot's (2000) recommendations for HEIs and health service partners to involve mentors and students in sharing the responsibility and contribute to developing assessment strategies to meet the needs of all those involved.

Finally, the theoretical frameworks identified in the papers reviewed in this IR largely centre on the use of taxonomies to determine levels of practice, despite criticism of their suitability, sensitivity and appropriateness to assess undergraduate nursing competence (Calman *et al.*, 2002). None of the tools and frameworks discussed seems to resolve the problems related to ambiguous language and accurate identification of performance levels. Therefore, a gap has been identified not only in the availability of appropriate tools to assist mentors and students in interpreting and differentiating between performance levels, but also in relation to a need for a suitable theoretical framework to underpin practice-based assessment. The next step was to consider what theoretical approaches might be appropriate to address the identified problems.

3.4 Development of the theoretical framework employed in this study

Having clarified the problems associated with the assessment process, the next element of the synthesising stage of this IR is to consider the theoretical underpinnings employed in practice-based competency development and assessment. From a wider theoretical standpoint, what emerges from the literature is that assessment does not have one overarching theory (Crisp *et al.*, 2003).

Consequently, a variety of theories from other fields such as decision-making and judgement theories seem to be adopted (Schuwirth and Van der Vleuten, 2011; Burden, 2014) and it is logical to assume the link to summative assessment. Formative assessment, on the other hand, although reviewed under the umbrella of assessment, is seen as an integral part of teaching and learning process rather than assessment, therefore it draws on general education theories (Gray, 1993; Hays and Wellard, 1998).

Govaerts and Vleuten (2013) discussed behaviourist, cognitivist, and constructivist learning theories, acknowledging their role in how well learning occurs, and viewed learning as a formal process that develops the learner from incompetent to competent. Similarly, Lave and Wenger (1991) refer to formative assessment as a 'community of practice', where a learner engages with workplace colleagues to acquire the necessary competencies to move from novice to expert. Working on the concept of situated learning, Lave and Wenger (1991) created the theory of disregarding the idea of assessor and learner in favour of 'full' and 'peripheral' participation. Learning, therefore, is by developing an identity of mastery, which occurs through participation in a 'community of practice'. To have learnt is to be able to participate more competently or more fully. Therefore, the interaction of the learner and assessor is an important consideration. Consequently, the theories underpinning assessment are expanded to include socio-cultural theories with their claim that learning emerges through active participation in a community, practicing in an authentic environment (Klenowski and Wyatt-Smith, 2014).

The notion that assessment is an integral part of the learning cycle, rather than being an episode that gauges students' performance at the end of a course, is gaining attraction by placing the emphasis on formative assessments. In this framework, assessment fits with the constructivist paradigm of teaching and learning (Elwood and Klenowski, 2002). Biggs's (2003) constructive alignment theory suggests that learning objectives, teaching and assessment should be aligned, explaining that if the achievement of learning is to be graded, there is a need for a constructively aligned, criteria-based assessment where grades are

awarded to specify the level of pass according to how well the desired learning outcomes are achieved. The key is that the teaching methods and the assessment tasks are aligned to the learning activities assumed in the intended outcomes (Biggs, 2003). Biggs and Tang (2007) described this in relation to 'surface' and 'deep' learning, with the latter instigating a conceptual change in students rather than merely the acquisition of knowledge.

In support of this argument, Schuwirth and Van der Vleuten (2011) pointed out that the central purpose of educational curricula is for learning to take place, so assessment should be aligned with this purpose (assessment *for* learning). Boud and Falchikov (2006) added that constructively aligned assessment impacts positively on future learning through providing targets that focus and drive deeper learning, whereas assessments that focus on the selection between the sufficiently and insufficiently competent students (assessment of learning) do not reach their full potential in steering student learning behaviour.

Assessment for learning

Assessment *for* learning emerged to challenge the traditional view that the purpose of assessment is to determine whether a student passes a test, regardless of whether enhanced learning has occurred (Popham, 2001; Ali, 2013; Haines *et al.*, 2013). The concept focuses on the formative potential of the assessment through individualised feedback on performance that is not designed for fail or pass. Instead, it provides a continuum where improvements are documented in individual areas as moving from 'working towards competence' to 'competent' thus refocusing attitudes towards assessment as something that can facilitate learning rather than a process that highlights incompetence. The aim in assessment *for* learning practices is to make students active participants in assessing their own learning.

Tillema *et al.* (2011) explained that assessment *for* learning scaffolds coherent, authentic, personalised and practical information to the student, helping them to actively engage in successful activities and facilitate reaching improved learning

outcomes. The hallmark of assessment *for* learning is to have assessment criteria that are clear, explicit, framed in language that is meaningful, and made available well in advance of the commencement of activities that will subsequently be assessed (Brown, 2005). Hence, assessment *for* learning is deemed to encompass constructive alignment (McDowell *et al.*, 2011).

Schuwirth and Van der Vleuten (2011) explained that assessment *for* learning is not a new theory, but more a change of views on assessment and much of the theoretical underpinning to support this approach still needs to be developed, despite complete consensus on how assessment influences learning, empirical research has not been performed in this area. The main strength of this concept is by making learning visible, supporting the development and achievement of performance and learning outcomes (Ali, 2013; Haines *et al.*, 2013).

Much of the theoretical rationale for the assessment *for* learning concept was provided by Black and William (1998), which was based on classroom assessment practices. They reported that the grading is over-emphasised in traditional assessments, leading to superficial learning and recall of isolated details that students soon forgot; additionally, when 'good marks' become the motivator, this encourages competition between students rather than personal improvement. Black and Wiliam (1998) proposed an assessment *for* learning as a process, characterised by shared responsibility between students and assessors to identify what students have to do in order to succeed, which increases students' self-concept and feeling of involvement and agency. The theory implies that assessors can bridge traditional boundaries using assessment *for* learning, as long as the assessment relationship is understood (Sadler, 1998).

Willis (2007, p. 52) argued that the concept of assessment *for* learning, involves students in routines of feedback and sharing of assessment criteria, which gives students motivation and autonomy to monitor their own learning progress, so they "become active meaning makers and thoughtful judges of their own learning". Students develop expertise through participating in the social interaction of working

with a more knowledgeable expert until they reach independence, shifting the focus of assessment towards mapping future learning growth and social support rather than measuring past performance (James, 2006).

Black and Wiliam (2006) observed that in assessment *for* learning practices, students can change their learning identity from that of passive recipients to being active learners who can take responsibility for their own learning. This is supported by Willis (2007) arguing that the goal of assessment *for* learning is to ensure that students develop identities as capable learners. A robust learning identity is judged by students' behaviour and skills that shows an increased participation in a community of practice.

Building a relationship between the student and the assessor that is based on trust and respect is fundamental in assessment *for* learning practices (Black and Wiliam, 2006). This encourages students to feel safe and engage with their peers or assessor in creating a shared language and understanding about the quality and meaning of performance. Therefore, according to Willis (2007), in assessment *for* learning, there is a paradigm shift from 'having power over students' to becoming one of 'sharing power with students'. Black and Wiliam (2006) acknowledge that such sharing is a significant challenge to the power and control that traditionally rests with assessors within an assessment relationship and explained that changes to the power balance when implementing assessment *for* learning practices can be exciting for some assessors and very threatening for others.

Within the context of practice-based assessment, Schuwirth and Van der Vleuten (2011) also argued that assessments focusing on the summative examination, although valuable in ensuring that students lacking the necessary competence do not graduate, do not provide information as to how an incompetent student can become a competent one, or how each student can strive for excellence.

McDermott *et al.* (2017) argued that in situations where students are assessed on their competence to perform in meaningful activity drawn from a professional context, non-traditional assessment methods that appraise the value of the

assessment process itself are required. Brown (2005) added that the design of assessment needs to be practice-orientated, emphasising the need to focus on assessment instruments that measure students' abilities to use the material they have learned in live situations.

The literature continues to support the view that grading practice needs to be more discriminating than the pass/fail system commonly used in assessing nursing students (Heaslip and Scammell, 2012; Burden, 2014; Terry *et al.*, 2017). The notion of assessment *for* learning, with clearly articulated expectations in criteria that describe levels of qualities in relation to each criterion, could inform a tool to help mentors not only to facilitate valid judgement of performance assessment but also to promote learning, self-assessment and improve feedback structure. There is a need, therefore, to develop an assessment process that is known to mentors and students in advance and communicates as clearly as possible what is required and what has been achieved.

Authentic assessment

Stemmed from the growing need for forms of assessment that can capture the complexities of teaching and learning as they develop overtime and across different contexts (Rennert-Areive, 2005), authentic assessment was first coined in 1989 by Grant Wiggins who called it "a true test" of intellectual achievement or ability. It requires students to demonstrate their deep understanding and complex problem-solving through performance of exemplary tasks that experts or professionals typically face in the field. Therefore, for an assessment to be authentic it needs to reflect the intellectual work of practicing professionals (Wiggins, 1989, p. 703).

Montgomery (2002) explained that authentic assessment involves assessing the application of process and product to the real world where problem-solving and critical thinking abilities are often used, and like assessment *for* learning, there is an alignment of curriculum content and assessment that allows learners and assessors to engage in meaningful learning, hence formative assessment can be easily incorporated into authentic assessments (Koh, 2017). Montgomery (2002) stressed

the importance of establishing the criteria before the period of instruction to focus the student and the instructor on the critical components of the curriculum, enabling them both to continue to modify the work in progress. As such, authentic assessment also captures students' dispositions such as positive habits of mind, growth mindset, persistence in solving complex problems and resilience (Koh, 2017).

McDermott *et al.* (2017) explored the pedagogical implications of authentic assessment, tracing it back to the pragmatic and constructivist philosopher John Dewey. Dewey (1897, cited in McDermott *et al.*, 2017) outlined a parallel view to authentic assessment commenting that the context in which learning takes place needs to prepare one for the real world and, as far as possible, should resemble real life.

Gulikers *et al.* (2004) explored authentic assessment in the context of professional education that incorporate competence-based assessments. They argued that to better prepare students for their future workplace, there is a need for assessment tasks used in professional education to resemble the tasks students will encounter in their future professional practice. Therefore, the application of authentic assessment can be used in a wider context, so it has value beyond the immediate learning situation. As discussed earlier, assessment of competence in clinical practice is equated with satisfactory performance of knowledge, skills and attitudes (Eraut, 1994; Yorke, 2005), which is directly tied to the 'real world', and hence competency-based assessment is considered a type of authentic assessment (McDermott *et al.*, 2017).

Scoring criteria used in authentic assessment should incorporate both the development of relevant professional competence and the assessment of students' learning progression (Koh, 2017). Therefore, in authentic assessment, the scoring criteria must be transparent and shared explicitly with students to facilitate their learning and judge their performance or work (Wiggins, 1989; Gulikers *et al.*, 2004; Koh, 2017).

In the literature, authentic assessment is frequently associated with the use of scoring rubrics, characterised by having clear and unambiguous descriptors that are applied consistently to compare students' performances against the descriptors (Wiggins, 1989; Montgomery, 2002; Gulikers *et al.*, 2004; Truemper, 2004; Frentsos, 2013). In a systematic review of grading practice within nursing, Gray and Donaldson (2009, p. 107) referred to scoring rubrics as "the most promising" method. Likewise, Heaslip and Scammell, (2012) advocated the use and testing of scoring rubrics. More recently, in their integrative literature review, Cockett and Jackson (2018) reported that the use of scoring rubrics has the advantage of enhancing feedback and self-assessment as well as the ability to understand the assessment criteria. The authors noted that, although the review set out to examine the implementation of scoring rubrics in nursing and midwifery, no papers were retrieved that are directly related to the disciplines. This indicates a gap in knowledge about the application of scoring rubrics in nursing which this research study will address. Therefore, a review of scoring rubrics is essential.

Scoring rubrics

A 'scoring rubric' is a matrix with concise description of expectations at different levels of accomplishment. These have been used extensively in academic settings but have not gained momentum in nursing, therefore, the benefits of scoring rubrics remain unrealised (Frentsos, 2013). A scoring rubric was defined by Montgomery (2002, p. 325) as "an assessment tool that uses clearly defined criteria and proficiency levels to gauge student achievement of those criteria. The criteria provide descriptions of each level of performance in terms of what students are able to do". There are two types of scoring rubrics: analytical (allowing for the separate evaluation of each component of the task that provides diagnostic information useful to both students and mentors) and holistic (generic type assessing overall quality). Scoring rubrics are made of three components: clearly defined performance criteria, detailed descriptions of what a performance looks like for each criterion, and a rating scale, commonly 3 or 4 points (Allen and Tanner, 2006).

What distinguishes scoring rubrics from other grading tools or checklists is the clearly defined performance criteria written in an easily understood language with descriptions of the possible level of attainment for each of the criteria or dimensions of performance. The criteria are described fully enough to make them useful for judgement of, or reflection on, progress towards valued objectives (Allen and Tanner, 2006). Thus, students and mentors have a clear idea of what constitutes excellence as the scoring rubric provides a clear summary for each level descriptor.

The association of scoring rubrics to the tenets of assessment *for* learning and authentic assessment is through employing 'real world' tasks that mutually engage students and mentors in meaningful activities, allowing for evaluation of effective performance (Wiggins, 1989). The transparency provided by employing the scoring rubric in advance, allows for identification of the student's areas of strength and weakness before the upcoming evaluation, encouraging both the student and the assessor to develop a collaborative and individualised action plan to reach the performance needed.

The advantages here are that mutual planning by involving the student in setting and attaining educational goals supports adult learning principles (Frentsos, 2013). In a study in the field of psychology, Halonen *et al.* (2003) used a scoring rubric as an authentic assessment strategy to facilitate and capture qualitative aspects of meaningful and enduring learning. They reported that the scoring rubric provided a framework to promote the relevance of course experience to students' lives and, as such, provided a road map for teaching effectiveness.

Additionally, the use of scoring rubrics has been suggested as a way to improve objectivity in assessment. It allows students to understand why they received a particular grade and provides formative assessment as both the assessor and the student know exactly what to do to achieve the higher level. This, in turn, facilitates individualised constructive feedback and self-assessment. Summative assessment is also achieved as the scoring rubric is used to award a final grade because the

criteria for meeting or not meeting the required level to pass is clearly established and articulated (Montgomery, 2002; Truemper, 2004; Frentsos, 2013).

In a key research review examining the reliability and validity of scoring rubrics, a total of 75 studies were appraised across many professional disciplines by Jonsson and Svingby (2007). They identified several benefits of using scoring rubrics in performance assessments. These included increased consistency of scoring, the facilitating of a valid judgement of complex competencies, and promotion of learning. The main conclusion was that reliable scoring of performance assessments can be enhanced through the use of scoring rubrics. They seem to have the potential for promoting learning; the main reason for this potential lies in the fact that scoring rubrics make expectations and criteria explicit and transparent, which also facilitates feedback and self-assessment. In addition, scoring rubrics facilitate communication and provide students, assessors and faculty with language to foster both feedback and discussions (Jonsson and Svingby, 2007).

Shipman *et al.* (2012) evaluated the positive and negative qualities of scoring rubrics, arguing that they provide a potential solution to the assessment by aiming to minimise inconsistencies in grading. They concluded that, despite some concerns such as assessor bias still being possible and that the promotion of conformity and standardisation in a manner that is incongruent with the concept of student-centred learning, a well-designed and formulated scoring rubric can be an effective assessment tool. Cockett and Jackson (2018) supported this claim, emphasising that engaging academic staff, students and practice assessors in co-creating and implementing scoring rubrics is essential for its success.

3.5 From the theoretical framework to study aims

The theoretical perspective discussed in this review highlights the need to establish a transparent and common language to interpret different levels of competence. Providing a structure that reduces confusion and ambiguity in interpreting competency statements may help mentors define what is expected of students and help students to identify what they are expected to achieve. Therefore, there is a

need to give both students and mentors a more effective system than is currently experienced.

Within the context of assessing the practice element of pre-registration nursing students, it is important to acknowledge that mentors' lack of understanding of competency statements or being able to make a decision on the appropriate level of competence, is essentially permitting people who have not met the required standards to enter the NMC register, creating a threat to patients' safety. This research study takes a pragmatic approach in seeking to improve the assessment of clinical competence. Adopting the concepts of 'assessment *for* learning' and 'authentic assessment', by having clearly articulated expectations in a scoring rubric that describes levels of qualities in relation to each criterion, may offer an effective and efficient means to enable mentors to differentiate between levels of competence more clearly. Therefore, it is reasonable to predict that this could help mentors not only to make valid judgements, but also could promote student learning by improving feedback and self-assessment, and subsequently, make practice-based assessment more fit for purpose.

In an effort to address the knowledge gap identified, and a clear and present need to make practice-based assessment of students fit for purpose, a potential solution identified in the previous research recommendations from studies included in this review is that a well-designed and transparent scoring rubric might address existing challenges (Donaldson and Gray, 2012; Heaslip and Scammell, 2012). Emerging from this review is a recognition that a collaborative approach that involves mentors and students in developing tools is essential for success in improving practice-based assessment (Norman *et al.*, 2002; Cassidy *et al.*, 2012; Cockett and Jackson, 2018). Additionally, as identified in the scoping exercise (see section 2.3) and frequently reported in the literature, mentors and students experience difficulty in identifying the required attitude and behaviour of aspirant members of a profession like nursing (Butler *et al.*, 2011; Strauss, 2016), hence such attributes are assessed more leniently as they are perceived more abstract and open to interpretation (Bondy,

1984; Hunt, 2014). Consequently, the design of a scoring rubric must involve mentors and students.

To reach agreement of various groups of stakeholders on a unified interpretation of level descriptors for the professional attitude, behaviour and responsibility statements within the PLPAD (Appendix 1), a consensus research method was deemed the most appropriate approach which would form the blueprint of a scoring rubric. Delphi consensus (as discussed in section 4.4.1) was chosen as the most suitable method. In addition, the usefulness of the designed scoring rubric needed testing by evaluating mentors' and students' experiences and perceptions on its impact. This evaluation needed to assess its effectiveness in providing mentors and students with a tool that facilitates interpreting and identifying different levels of competence, therefore, minimising subjectivity of judgements. Additionally, evaluating the scoring rubric effectiveness in facilitating constructive feedback and self-assessment was also needed. Ultimately, the aim was to evaluate if a consensus-based scoring rubric makes practice-based assessment more fit for purpose. These considerations informed the development of the study research aim and questions.

3.6 Research aim, questions and objectives

Research aim

The aim of this study was to make practice-based assessment of pre-registration nursing students fit for purpose.

Research questions

The study had the following sequential research questions:

1. Can consensus be achieved among stakeholders on how to interpret level descriptors for the professional attitude, behaviour and responsibility statements in the PLPAD?
2. Does a consensus-based scoring rubric improve interpretation of level descriptors?

3. Does a consensus-based scoring rubric strengthen the rigour of mentors' assessment?
4. Does a consensus-based scoring rubric enhance learning, self-assessment and feedback provision?

Research objectives

- To establish, through a Delphi study, the degree of consensus or divergence between stakeholders in interpreting level descriptors for the professional attitude, behaviour and responsibility statements in the PLPAD.
- To develop a consensus-based scoring rubric using the findings from the Delphi study.
- To evaluate the effectiveness of the consensus-based scoring rubric for students and mentors by evaluating its effectiveness in:
 - Providing clear language and level descriptors.
 - Strengthening the rigour of mentors' assessment.
 - Enhancing learning, self-assessment and feedback provision.

3.7 Chapter summary

This chapter presented the integrative review that systematically synthesised empirical and theoretical literature, establishing that mentors and students struggle to understand the language and level descriptors used in practice assessment documents, which is likely to result in invalid and unreliable assessment of competence. The chapter identified that a well-designed scoring rubric with a transparent and common language to interpret different levels of competence might offer the solution to the challenges faced in practice-based assessment by helping mentors define what is expected of students and for students to identify what they are expected to achieve. The chapter concluded with the research aim, questions and objectives. The next chapter will discuss the philosophical and methodological approaches taken in the study with an overview of the methods employed.

Chapter 4: Ontology, Epistemology, Methodology and Methods

4.1 Introduction

The previous chapter identified the gap relating to ambiguity of the terminology and levels of competence in practice-based assessment documents. Based on empirical and theoretical evidence, designing and evaluating a scoring rubric may offer a solution. This chapter outlines the overarching philosophy that guides the epistemological and methodological approaches for this study.

The chapter comprises three sections. The first section discusses the philosophical and epistemological approach of this research: critical pragmatism. It commences with an overview of key features of classic pragmatism before exploring critical pragmatism espoused by John Dewey (1859–1952). The second section discusses the methodology applied in this research, mixed methods, justifying why this was favoured as the most appropriate and pragmatic approach for answering the research questions. The third section explores the methods used in this research (Delphi and evaluation research) including links to both epistemology and methodology.

Ontological and epistemological position

Before embarking on this research study, I was not sure about my ontological and epistemological views of the world and about the nature of reality and knowledge. I know I am practical and direct to the point and my approach to problems is trying to find solutions and what I teach my children is that 'if you can't solve it, find out how to work around it'. I am a strong believer that trial and error is an effective strategy and should be encouraged. This is reflected in my approach in this study by trying to search for and try solutions to the problems associated with mentors not accurately assessing nursing students' clinical competence. I recognised this in myself before I knew what pragmatism is. Then I found that pragmatism is all about being practical, looking for solutions and if it does not work look for another way to work it.

In my experience, as a nurse and as an educator, I realised that there is no such thing as absolute truth; it is highly likely that, despite what is practiced today as based on evidence, new knowledge frequently comes to prove that is not the case. This tentative approach to knowledge influences not only how I view the world but also my approach when I teach my students to help them function in the unpredictable real world and be open-minded and seek to keep themselves up to date. Pragmatism explains elegantly how I see the world; that the theory of evolution includes shaping our mental faculties and that knowledge changes as we interact with our peers and the environment.

4.2 Ontology and epistemology

When initially considering the philosophical approach for this project, it was clear that the objectives of the study, comprising quantitative and qualitative approaches in all stages, did not fit either the positivist or interpretive paradigm exclusively. Involving stakeholders to agree on language used to interpret the competencies arguably may align with critical social theory, namely communicative action and action research.

Communicative action (Habermas, 1984) assumes that language and meaning are the foundational component of the coordination of social action to achieve mutual understanding and '*truth*'. Within this paradigm, research becomes a means for 'taking action' by explaining how things could be, emphasising the process not the product, therefore emancipation becomes the goal (Maguire, 1987). It was decided that communicative action would only partially achieve the study objectives since, although it could apply to achieving mutual interpretation of the statements needed to design the scoring rubric, it would not align naturally to the experimental stage of evaluating the effectiveness of it.

Another approach drawing connections to critical social theory is participatory action research. This approach was also considered, owing to an identified need for stakeholders' participation in designing the scoring rubric to enhance the assessment process in practice (Friere, 1970). Action research lends itself to the

epistemological rationale that knowledge is embedded in social relationships and is most influential when produced collaboratively through action (Hawkins, 2008). In this approach, rather than doing research 'on' others, the researcher is considered an 'insider' who is part of the culture and is involved in cycles of interventions during data collection (Bresler, 1995). Although key features of action research met the objectives of this study, it was not considered suitable due to the researcher's intention to maintain the traditional 'outsider observer' to ensure objectivity and eliminate personal perspectives and bias from influencing the interpretations of the competency statements. This separation was important given the continuing criticism that existing assessments are frequently subjective.

In their review of paradigms used in nursing research, Weaver and Olson (2006) concluded that the effectiveness of nursing inquiry should be measured by its problem-solving ability or usefulness to those involved. Pragmatic approaches aligned well with the objectives in this study, which involved taking practical steps to design a scoring rubric in an attempt to solve the inconsistencies in the way competency statements are interpreted and the difficulties in identifying different levels of competence. In particular, John Dewey's (1859–1952) contribution to pragmatism, that fosters cooperative and social inquiry aligns well with developing agreement over how competency statements should be expressed and assessed.

Pragmatism, derived from the Greek word 'pragma' meaning action, is a philosophical tradition that originated in the United States in the 19th century. Founded by Charles Sanders Peirce (1839–1914), for him, pragmatism was primarily a philosophy of meaning; for any statement to be meaningful, it must have practical bearings and application in the real world. Peirce stated in his classic paper *the fixation of belief* (Peirce, 1877) that inquiry is a struggle to replace doubt with settled belief, arguing that an experimental method is the only method that can make sense of the fact that we are disrupted by inconsistent beliefs. This broadly places pragmatism in the realist tradition. However, it was William James (1842–1910) who gave it the name 'pragmatism' in 1907 with his book *Pragmatism: A New Name for Some Old Ways of Thinking*. James built on Peirce's idea that 'truth' is

what reality compels individuals to believe in a given context, and that its acquisition is through extermination rather than a revelation, rejecting the idealist view that reality and knowledge are determined by the structure of human thought (Guba and Lincoln, 2005).

Pragmatists were captivated by Charles Darwin's publication of *The Origin of Species by Means of Natural Selection* (1859). It was not so much the theory of evolution that pragmatists found interesting, but rather the process of natural selection itself that is based on "chance errors leading to uncertainty and fallibility" (Menand, 1997, p. 210). Pragmatists believe that evolution is not just a biological *truth*, drawing attention to the impact of evolution in shaping our mental faculties (James, 1890). They believe that a person's knowledge and being is continuously changing and shifting and that the universe is in a constant changing process. Because everything in the world, including knowledge, is perpetually changing, interaction with peer and environment is fundamental to pragmatism. This interaction involves linking processes to consequences and comparing them to desired ends such as improving the world or aspects of it (Wolfe, 2012).

While they are drawn towards the realists in employing scientific methods to establish what is '*truth*', pragmatists are open-minded, tentative and modest about the certainty of absolute *truth* since it is subjective and relative to our consistently evolving desires of "what works". Such an anti-foundational dimension of pragmatism sharply diverges from the realist claim that the world exists apart from our understanding of it and the idealist assumptions that the world is created by our conception of it (Shields, 1998; Guba and Lincoln, 2005).

John Dewey (1859–1952) was the heir to the pragmatic tradition from William James and Charles Peirce. He continued to develop pragmatism and focussed its application onto practical issues such as education. In his book, *Democracy and Education* (1916/1966), Dewey established that education philosophy is consistent with the pragmatic idea and embraced the theory of natural selection. The way our mind works relates to our interaction and adaptation to the environment, and

knowledge is the result of an adaptive human response to events or processes of change. To make more sense of the world, Dewey explained that ideas are used as instruments of inquiry that require active participation in processes to reconstruct events with alternative variations, and, concluded that testing the practical instrumentality of the proposed change must be carried out experimentally.

Ideas are anticipations of possible solutions. They are anticipation of some continuity or connection of an activity and a consequence, which has not as yet shown itself. They are therefore tested by the operation of acting upon them. They are to guide and organise further observations, recollections, and experiments (Dewey, 1966, p. 160).

Acknowledging that experimentation itself is fallible, Dewey (1966, p. 154) said that “trial and error is very powerful”. Errors would mitigate further and more rigorous inquiry leading to higher levels of thinking processes. The act of thinking or trials and errors of an experience allows learners to think about their actions and their consequences.

Dewey saw inquiry as a democratic process that places the focus on practicality and social involvement in deciding what is *truth* and *knowledge*. In Dewey’s view, this fosters a cooperative and social inquiry where different groups can learn how to contribute to their democratic society. The purpose is to produce citizens who are trained how to make choices in what Dewey called ‘associated living’, doing things with other people, demanding that one must experience another person’s opinions and beliefs in order to more fully develop their own (Dewey, 1916/1966). This view aligned well with the objectives in this study where different groups of stakeholders involved in pre-registration nursing education participated in a social inquiry to reach common language for the professional attitude, behaviour and responsibility competency statements in the PLPAD.

This dimension of Dewey’s pragmatism has been branded with the contemporary name ‘critical pragmatism’ (Kadlec, 2007). Critical pragmatism emerged as an

approach that continues to hold the pragmatic principles of uncertainty about absolute *truth*, and is inspired by Dewey's idea of knowledge as a critical form of inquiry to find a practical solution that values social action to serve the public good, with emphasis on involvement, participation and openness as a tool to change (Kadlec, 2007). In the case of this study, the public good that needs to be served is to have more robust assessment of clinical competence to ensure unsafe practitioners do not achieve registration as nurses.

A number of authors report the origins of the term "critical pragmatism". Deegan (1988) associated it with the pioneering social reformers Jane Addams (1902, 1910) and Dewey (1909, 1916) combining pragmatic philosophy with democratic and social values, thereby bringing critical theory and pragmatism closer together (Ulrich, 2007). Kadlec (2006, 2007) associated critical pragmatism with Habermas's critical social theory, reconstructing Dewey's pragmatic thoughts to appreciate its "critical" features. White (2004) also related pragmatism to critical theory, claiming that to redeem the unfulfilled promise of providing a systematic orientation to critical social science, critical social theory needs to be revised to develop a "pragmatically configured critical theory" (White, 2004, p. 311).

However, both traditions continue to differ depending on whether their point of departure is primarily critical theory or pragmatist philosophy and whether their aims are more theory or practice oriented (Ulrich, 2007). Critical pragmatism provides an interesting perspective that is relevant to nursing in bringing theory and practice closer together. A theory-practice gap exists when theory does not address practice demand; a pragmatic approach calls for theory to be designed and tested in practice (Weaver and Olson, 2006).

Critical pragmatism was found to be the most suited overarching ontological and epistemological approach for this study. It aligned with the approach taken to design the scoring rubric by employing an alternative that ensured social involvement in deciding what is *truth* and *knowledge* when interpreting the competency statements in PLPAD. This was followed by testing the new approach

to evaluate the “practical instrumentality” of the proposed change (Dewey, 1966, p. 160).

4.3 Methodology

Methodology has been described as “a broad approach to scientific inquiry specifying how research questions should be asked and answered” (Teddlie and Tashakkori, 2009, p. 21), and provides the theoretical perspective that links a research problem with a particular method or methods (Hesse-Biber, 2010).

When considering the most appropriate methodological approach to answer the research questions in this study (see section 3.6), the methodology had to successfully address all aspects of the research. This includes the ontological and epistemological underpinning, data collection, analysis and inference techniques (Johnson *et al.*, 2007). Interpreting and developing level descriptors for the competency statements in the PLPAD by stakeholders involved in pre-registration nursing education indicated that a qualitative approach would provide the best fit. An interpretive qualitative methods approach, however, would be of questionable suitability for measuring the level of agreement on the generated interpretations of statements.

Additionally, to evaluate if mentors and students found the scoring rubric effective, collecting both quantitative and qualitative data would be necessary to confirm findings. Therefore, it was recognised that answering the research questions could not be achieved through the application of solely quantitative or qualitative methodologies, so a pragmatic strategy of mixing more than one research method was chosen on the premise that the research questions dictate the type of research design (Creswell and Plano Clark, 2017).

4.3.1 Mixed methods research

A mixed methods approach is considered appropriate where one approach alone is insufficient to answer the research question. Historically, quantitative and qualitative methods were seen as intrinsically different separated by a 'paradigm war' based on their different ontological and epistemological paradigms (Johnson and Onwuegbuzie, 2004; Guba and Lincoln, 2005). However, since the 1960s, there has been an acceptance of different paradigms where researchers have increasingly challenged the dominance of a single method resulting in the emergence of mixed methods as a third paradigm (Tashakkori and Teddlie, 1998). Several definitions for mixed methods have emerged over the years that incorporate various elements of research methods, processes, purpose and philosophy (Creswell *et al.*, 2017). Johnson *et al.* (2007) sought a consensus on a definition based on understanding different definitions provided by highly published mixed methods researchers, concluding with their composite definition:

Mixed methods research is the type of research in which a researcher or a team of researchers combined elements of qualitative and quantitative research approaches for the purpose of breadth and depth of understanding and corroboration (Johnson et al., 2007, p. 123)

Creswell and Plano Clark (2017) criticised Johnson's (2007) definition for not specifically mentioning paradigms or philosophies, arguing that the definition of mixed methods continues to be contested. In their latest book edition, Creswell and Plano Clark (2017) were inclined to provide a definition of core characteristics of mixed methods research that combines methods, research design and philosophy orientation. Interestingly, they emphasised the researcher as the key component that goes into designing and conducting a mixed method study. Figure 4.1 (overleaf) outlines Creswell and Plano Clark's (2017) definition of core characteristics.

The researcher

- Collects and analysis both quantitative and qualitative data rigorously in response to research questions and hypotheses.
- Integrates the forms of data in their results.
- Organises these procedures into specific research design that provides the logic and procedures for conducting the study.
- Frames these procedures within theory and philosophy.

Figure 4.1: The definition of core characteristics of mixed methods research (Adapted from Creswell and Plano Clark, 2017).

Mixed methods continue to evolve and develop as a viable alternative that can be included in all phases of research (Johnson and Onwuegbuzie, 2004). It is described as a methodology with orientation towards social inquiry that actively invites dialogue about multiple ways of making sense of the social world and multiple standpoints on what is important to be valued and cherished (Greene, 2008). Consequently, mixed methods is a way of thinking that rests on the assumption that the complex character of social phenomena requires the use of multiple approaches, and that any given approach to social inquiry is inevitably partial (Greene, 2008). In other words, mixed methods research is a philosophy that takes a pragmatic position by attempting to fit together the insight provided by quantitative and qualitative research mixed in a way that offers the best opportunities for answering important research questions (Johnson and Onwuegbuzie, 2004).

Therefore, mixed methods ensure clear alignment between the research questions, design and the pragmatists' principles articulated by Dewey (1948, p. 132) that methods are selected for their "practical consequences and empirical findings" to decide which action to take next to understand the 'real world' phenomena and determine future action. It frees the researcher from the constraints of following one particular tradition (Brannen, 2005).

The decision to use mixed methods in this study was based on the research questions that focused on interrelated and multidimensional stages. Eliciting the

stakeholders' interpretations of the competency statements in the PLPAD was best sought using a qualitative method to provide them with a platform to share their own opinions. Quantitatively calculating consensus between participants was needed to measure the level of agreement as well as providing them with a broader insight into where their individual reality fits among others.

The designed scoring rubric was then experimentally piloted. Both quantitative and qualitative methods were used to evaluate the scoring rubric's impact on the assessment process in the practice settings. Integrating quantitative and qualitative data collection and analysis proved to be methodologically robust as the findings from the qualitative element enhanced interpretation of the quantitative results, to further support the selection of the mixed methods approach to facilitate enhancement of findings with a second method (Creswell and Plano Clark, 2017).

Consideration was also given to the ordering of collecting and analysing quantitative and qualitative data; however, in discussing mixed methods research, it is important not to neglect the fact that methods may be combined within either the quantitative or qualitative paradigm (Stange *et al.*, 1994). When mixing methods, Creswell and Plano Clark (2017, p. 59) emphasise the importance of addressing the "point of interface", where the mixing or integration occurs.

To guide researchers in selecting logical research methods that ensure the resulting design is rigorous, Creswell and Plano Clark (2017) outlined three core mixed methods designs. These comprise convergent design (results of quantitative and qualitative data analysis compared or combined concurrently), explanatory sequential design (starts with the collection and analysis of quantitative data), and exploratory sequential design (starts with the collection and analysis of qualitative data). Figure 4.2 (overleaf) outlines the latest Creswell and Plano Clark's (2017) three core designs⁶.

⁶ In their most recent third edition in 2017, Creswell and Plano Clark changed reduced the MM designs to three core designs (presented in Figure 4.2), compared to the seven in the 2011 edition. However, the two designs used in this study remain the same

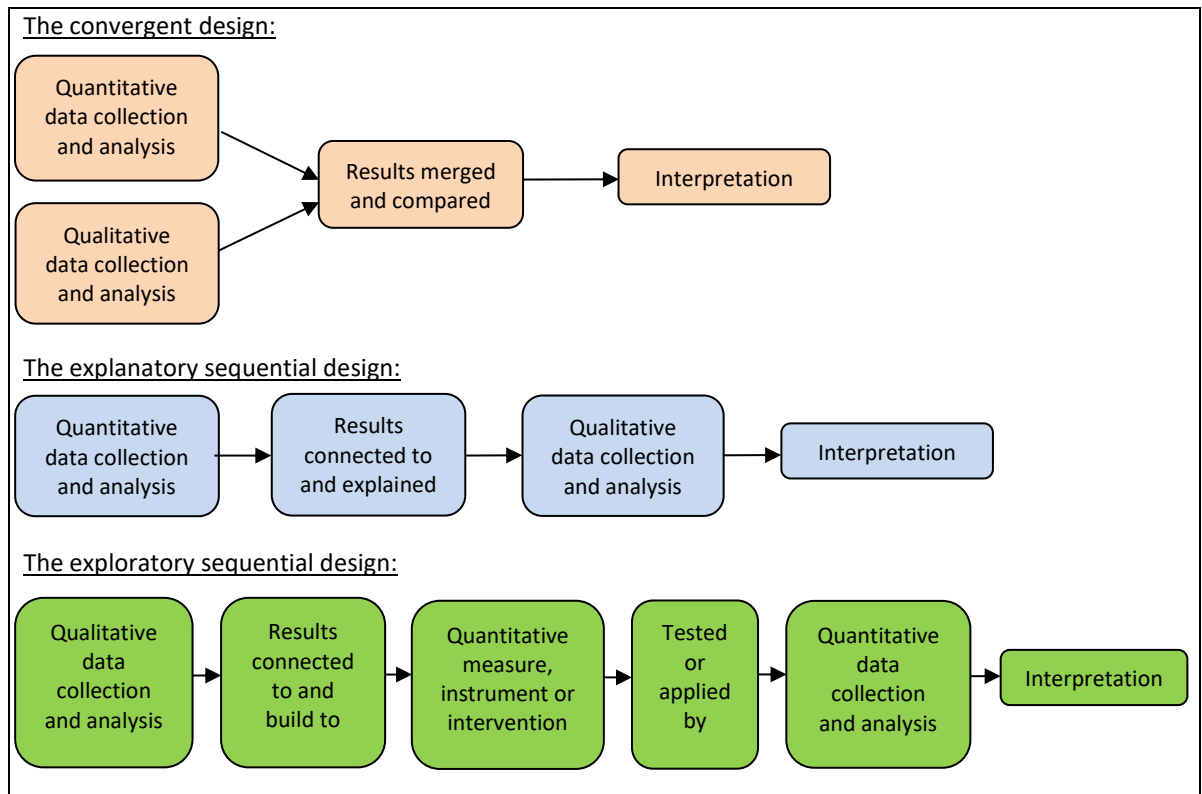


Figure 4.2: The three core mixed methods designs (Adapted from Creswell and Plano Clark, 2017)

The order of mixed methods design used in this research study was influenced by the nature of the research objectives. The Delphi phase in this study adopted an ‘exploratory sequential design’. In this design, the researcher starts with collecting and analysing qualitative data that is then followed by a developmental stage of translating the qualitative findings into an approach or a tool that is tested quantitatively (Creswell and Plano Clark, 2017). This provided a clear justification that adopting exploratory sequential design for the Delphi phase of this study was appropriate, since the aim was to generate statements to design the scoring rubric through a qualitative method (round one of Delphi), and build on them sequentially by a quantitative method to establish consensus (round two of Delphi).

The evaluation research phase adopted a convergent design, in which the researcher collected quantitative and qualitative data at the same time, and then merged the two databases for the purpose of comparing or combining the results (Creswell and Plano Clark, 2017). The convergent design was suitable for this phase of this study since the aim was to evaluate the effectiveness of the scoring rubric

through collecting quantitative and qualitative data concurrently, then integrate the quantitative results and qualitative findings for the purpose of complementing each other to reach a fuller and *authentic* answer to the research questions. Figure 4.3 outlines the different mixed methods designs used in this study.

Phase one: Development of the scoring rubric				
Research method	Mixed methods design	Tool	Application	Timeline
Consensus Delphi	Exploratory sequential design	Questionnaire (round 1)	Qualitative data collection and analysis of stakeholders' interpretations.	7 th September 2015 to 12 th October 2015
		Questionnaire (round 2)	Quantitative data collection and analysis of level of consensus.	10 th November 2015 to 6 th December 2015

↓

Phase two: Evaluating the impact of the scoring rubric				
Research method	Mixed methods design	Tool	Application	Timeline
Evaluation research	Convergent design	Two questionnaires (a student and a mentor version)	Concurrent quantitative (Likert scale) and qualitative (free-text) data collection and analysis.	14 th November 2016 to 17 th February 2017

Figure 4.3: The mixed methods designs used in the study.

4.3.2 Mixed methods validity

The development of mixed methods has not been without controversy. Besides the paradigm war debate, Onwuegbuzie and Johnson (2006) explained that in comparison to establishing validity in quantitative research and trustworthiness in qualitative research, discussions about validity issues in mixed methods research are in their infancy and assessing the validity of integrated findings is particularly complex.

In the literature, early discussions around threats to validity in mixed methods research focused on identifying both quantitative and qualitative approaches separately (Creswell and Plano Clark, 2017). Tashakkori and Teddlie (2003) examined validity in mixed methods research and contend that threats may occur from two aspects of the research: design quality (adequacy of the use and

implementation of the process used to reach conclusions), and interpretive rigour (the consistency of the conclusions with other aspects of the research).

According to Creswell and Plano Clark (2017), validity in mixed methods research requires employing strategies that address potential issues in data collection, data analysis and the interpretations of emerging combined results. Several scholars have discussed criteria to evaluate threats to validity in mixed methods research (Onwuegbuzie and Johnson, 2006; Dellinger and Leech, 2007; Tashakkoori and Teddlie, 2008; O’Cathain, 2010), however, the integrative framework (Tashakkoori and Teddlie, 2008) and the legitimization framework (Onwuegbuzie and Johnson, 2006) are frequently suggested (Ihantola and Kihn, 2011).

Ihantola and Kihn (2011) reviewed the criteria used to evaluate threats to validity in mixed methods research and concluded that the integrative framework may give a false impression that validation is an outcome only. They concluded that Onwuegbuzie and Johnson’s (2006) legitimization framework is more extensive, introducing several novel dimensions to the validity of mixed methods research not addressed before, therefore, the legitimization framework appears to be the most promising in addressing validity threats specific for mixed methods research (Ihantola and Kihn, 2011).

Onwuegbuzie and Johnson (2006) explained that ‘legitimation’ is not an outcome, but a continuous, iterative and interactive process that deals with threats to internal and external validity or credibility in quantitative and qualitative research respectively. Hence, it should occur at each stage of the mixed methods research process, whether quantitative, qualitative or both. Onwuegbuzie and Johnson (2006, p. 57) described nine types of legitimization threats that result from combining inferences from the quantitative and qualitative components of a mixed methods research study to form ‘meta-inferences’. Table 4.1 (overleaf) present the nine types of legitimization and how they were addressed in this research study.

Table 4.1: Legitimation framework adapted from Onwuegbuzie and Johnson (2006) and how it was addressed in this research study.		
Legitimation type	Description	Addressed in this study
Sample integration	The extent to which the relationship between the quantitative and qualitative sampling designs yields quality meta-inferences.	Same individuals (and groups) were involved in both mixed methods of the research.
Inside-outside	The extent to which the researcher accurately presents and appropriately utilises the insider's view and the observer's views for purposes such as description and explanation.	The researcher maintained the objective outsider view in all phases of the research and was challenged for subjectivity by experienced researchers (the supervisors).
Weakness minimisation	The extent to which the weakness from one approach is compensated by the strengths from the other approach.	The two-phase mixed methods allowed findings from one method to be used to complement the findings from the other, so that possible threats and weaknesses from one method can be compensated by the strengths of the other.
Sequential	The extent to which one has minimised the potential problem wherein the meta-inferences could be affected by reversing the sequence of the quantitative and qualitative phases.	The sequential design was used for the Delphi rounds and reversal of sequences was not possible. However, validity was addressed within the method (see section 5.4.3).
Conversion	The extent to which the quantitising or qualitisng yields quality meta-inferences.	Numbers were used to enhance the narrative; numbers and percentages were used to describe small samples, so they were not misleading.
Paradigmatic mixing	The extent to which the researcher's epistemological, ontological and methodological beliefs that underlie the quantitative and qualitative approaches are successfully (a) combined or (b) blended into a usable package.	Adopting a critical pragmatism approach was suitable for separating and combining research phases to answer the research question.
Commensurability	The extent to which the meta-inferences made reflect a mixed worldview based on the cognitive process of Gestalt switching and integration.	The process to reach inferences involved constant discussions with the supervisors to encourage back and forth Gestalt switches between a qualitative lens and a quantitative lens to ensure richer perspectives.
	The extent to which addressing legitimation of the quantitative and qualitative components of the study result from the use of quantitative, qualitative, and mixed validity types, yielding high quality meta-inferences.	The research addressed quantitative validity and qualitative trustworthiness in the appropriate parts of the study, (see sections 5.4.3 and 6.4.2).
Political	The extent to which the consumers of mixed methods research value the meta-inferences stemming from both the quantitative and qualitative components of a study.	In line with critical pragmatism, the research aimed to answer an important research question and introduced a workable solution to a problem. This advocates that mixing methods produced practical results that would be valued by research consumers.

In this study, evaluation of methodological rigour was achieved by employing established and relevant processes for both the quantitative and qualitative approaches. Additionally, the findings from the quantitative and qualitative components were integrated to combine inferences stemming from both methods to generate new perspectives and insight to create a bigger and fuller picture, drawing on the strength of both approaches to stimulate an inquiry that complemented one paradigm with another (Onwuegbuzie and Johnson, 2006). Brannen (2005) argued that this is an important advantage of using mixed methods.

4.4 Methods

As discussed in the previous section, mixed methods were used for the two phases of data collection. This section outlines the methods used in each phase of this study to answer the research questions. The first section provides an overview of the Delphi method, chosen as the most suitable method to generate stakeholders' consensus on how the professional attitude, behaviour and responsibility competency statements in the PLPAD should be interpreted. The second section reviews the evaluation research method used to evaluate the impact of the scoring rubric on the assessment of student nurses in the practice settings.

4.4.1 Consensus methods

In keeping with the theoretical underpinning of assessment *for* learning and authentic assessment, it was crucial to select a suitable method that ensures the interpretation of level descriptors for the competency statements are unambiguous, free from academic jargon and reflect the complexity of real life. Equally, the method should reflect critical pragmatism principles of the involvement and participation of society in a critical inquiry to introduce change (Dewey, 1966; Kadlec, 2007). Consensus methods were found to be well-suited in providing the means of encapsulating the insights of different stakeholders, to gain an authentic interpretation and collaborative group decision-making aimed at reaching the agreement of participants (Burgess and Spangler, 2003).

Consensus methods are increasingly being used in healthcare research and are acknowledged as effective approaches (Cantrill *et al.*, 1996). The main premise of consensus methods is based on the assumption that the opinion of a group is of more value than that of an individual (Keeney *et al.*, 2011). In deciding the most appropriate method, the three commonly used consensus methods in health research were reviewed: nominal group technique, consensus conference, and Delphi method.

The nominal group technique uses a structured group approach led by a moderator and gives participants an equal voice as they write their answers individually and silently, then the answers can be shared. The process can be repeated several times with the aim of reaching a consensus on a set of prioritised solutions representing group preferences (Carney *et al.*, 1996). The consensus conference is organised by inviting a group to a conference venue, then following a presentation or discussion around the issue at hand, delegates can vote on their preference or decision on the issue (Jones and Hunter, 1995).

Although nominal group technique and consensus conference have the advantage of face to face discussions that allow better understanding of the issues requiring consensus in a relatively short amount of time, they present the logistical difficulty of arranging for all participants to meet at the same time and risk of dominant individuals or groups dictating the direction of the discussion (Keeney *et al.*, 2011). In the same way, face to face meetings may increase the influence of social desirability bias, where members may not disclose their true perspectives if they feel they might be judged as professionally or socially inappropriate (Ecken *et al.*, 2011).

In contrast with nominal group technique and consensus conference, Delphi consensus is a method used to assemble and refine judgements of a group of experts (called a panel) using a series of questionnaires without the need to meet face to face (Linstone and Turoff, 1975; Powell, 2003). This method offers the advantage of allowing larger groups of people to take part across many

geographical locations at their own convenience (Keeney *et al.*, 2011). In addition, it eliminates the risk of any one participant dominating or influencing the group, avoiding potentially destructive group dynamics (Jairath and Weinstein, 1994).

Guided by the nature of the research objectives and the study needs, the practical and logistical characteristics of the Delphi method made it more attractive than any other consensus method for the purpose of data collection from a diverse group of stakeholders who are geographically dispersed across the UK. In addition, this method would add rigour to this study since the invited stakeholders (people invested or involved in pre-registration nurse education) have different power and influence bases.

4.4.2 The Delphi method

In ancient Greek mythology, the name 'Delphi' is derived from the Oracle of Delphi, one of the most important Oracles, who was believed to have the ability to interpret and predict the future (Linstone, 1978). People from far and wide travelled to Delphi to consult the Oracle on a range of important public or private matters, making the term 'Delphi' synonymous with receiving good judgement on an issue (Keeney *et al.*, 2011).

The Delphi consensus method took an academic form in the 1950s, when the RAND Corporation in the United States introduced the technique as a scientific study to obtain intuitive insight and judgement of experts on future trends in military defence projects (Linstone and Turoff, 2011). Since then, the Delphi method has gradually found its place in academic studies, becoming highly popular in the early 21st century (Landeta, 2006). It is described as "a method used to systematically combine expert knowledge and opinion to arrive at an informed group consensus on a complex problem" (Donohoe and Needham, 2009, p. 416). In addition to its forecasting features, Delphi consensus method is useful in dealing with issues that can benefit from collective judgement and, as such, has been advocated for identifying components of professional effectiveness and core skills to be evaluated

on clinical placements (Hicks, 2009; Lock, 2011; Penciner *et al.*, 2011; Boerner *et al.*, 2015).

The term 'Delphi' is frequently used in the literature interchangeably with 'Delphi method' (Linston and Turoff, 1975), 'Delphi study' (Cooney *et al.*, 1995), 'Delphi technique' (Hsu and Sandford, 2007), 'Delphi survey' (Hasson *et al.*, 2000) and 'Delphi process' (Buck *et al.*, 1993) to have the same meaning. For consistency, the term 'Delphi' is used throughout this thesis.

Criticism of Delphi has been reported in the literature. Much of the early criticism was centred on poor research techniques ranging from the questionnaires being poorly worded and ambiguous (Hill and Fowles, 1975) to superficial analysis of responses (Linstone and Turoff, 1975). Rowe and Write (2001) argued that, although recent research has generally been conducted by social scientists using standard experimental procedures, little evidence has accumulated regarding how best to conduct Delphi and when to use it. They stressed that to truly understand Delphi, controlled studies are needed to focus on what makes Delphi work such as influences of feedback, panel compositions and sizes, and how panellists' judgements change. Nevertheless, Rowe and Write (2001) emphasised that Delphi was designed for use with experts, particularly in cases where individual panellists have only limited knowledge and might benefit from communicating with others possessing different information.

In relation to the objective and technique of Delphi, Linstone and Turoff (1975) characterised Delphi as a method for structuring a group communication process, so that the process is effective in allowing a group of individuals as a whole to deal with complex problems. Rowe and Write (2001) expanded on the Delphi characteristics by identifying four features: anonymity, iteration, controlled feedback, and statistical aggregation. Table 4.2 (overleaf) demonstrates how these characteristics were addressed in this study.

Table 4.2: Delphi characteristics (Rowe and Write, 2001) and how they were addressed in this study.		
Characteristic	Description	Addressed in this study
Anonymity	Participation privately through using self-administered questionnaires allows free expression of opinions without social pressures from other members of the group.	The questionnaires were administered through a secured online host (Bristol Online Survey) using a free-text first round. Participants were assured that their responses were anonymous.
Iteration	Facilitated through a series of questionnaires allowing participants to refine or reconsider their opinions from round to round based on the opinions and judgements of all group members and not just the most vocal or powerful.	The responses from the first round were analysed and sent back to the participants to review or reconsider their own opinions. All participants have the same platform, preventing dominant individuals influencing opinions.
Controlled feedback	The group members are informed of the opinions of their anonymous colleagues.	Participants were provided with a summary of all opinions.
Statistical aggregation of responses	Facilitates quantitative analysis of responses, usually comprising a statistical average of mean or median values.	Attitudinal scale was used to collect responses from the second round to determine level of consensus.

Having the opportunity to identify and agree the components, the consensus reached among the group can be considered as achieving face, content and concurrent validity (Williams and Webb, 1994). However, it is important to note that the purpose of Delphi is to elicit opinions which includes gathering different perspectives. Divergence of opinions is equally useful in acknowledging the complexities of reaching agreement (Keeney *et al.*, 2011).

The original Delphi (classic Delphi) entails two or more rounds of questionnaires with a qualitative first round to enable the identification of a wide array of views (Linstone, 1978). The data collated and analysed by the facilitator is then used to generate statements in a questionnaire and then redistributed to the panel for consideration (round two). The panel is asked to rank their degree of agreement or reconsider their opinion when they see the views of others. This process is repeated until consensus is reached or when no further changes are taking place (Keeney *et al.*, 2011). The process in this study was terminated after two rounds when a pre-determined consensus for all items was reached (see section 5.6.2).

Traditionally, the questionnaires in classic Delphi are conducted by sending out a standard letter, but with technological advances, “e-Delphi” is commonly used to conduct the study online by using emails for communication and data collection (Avery *et al.*, 2005). The ‘classic e-Delphi’ approach was used in this study (see section 5.2).

There are several advantages of using e-Delphi including the convenience for participants to select when to participate. Donohoe and Needham (2008) claimed that using the internet to communicate and administer the questionnaires may reduce attrition by shortening the time between rounds. Cost saving is another advantage as emails and reminder emails are sent out without the need for postage and packaging (Keeney *et al.*, 2011). Data management compared to the paper-based method is also more convenient as computerised programmes communicate with each other; hence there is no need for transcription, and results can be easily transferred to other databases (Donohoe and Needham, 2008). Participants with special learning difficulties such as dyslexia may wish to spend longer time or control the screen size and background colour to suit their needs, which is considered as a benefit compared to the printed questionnaires.

On the other hand, with technological advances new problems could emerge including some participants not having email accounts or having limited physical or technical abilities especially when service users are being invited to participate. Other challenges include computer firewall settings blocking emails or directing them to quarantine or Junk folders (Keeney *et al.*, 2011).

There is no definitive answer to what is the optimal number of rounds required in a Delphi. Rohrbaugh (1979) and Rowe and Wright (1996) showed accuracy improvement over two to three rounds and Erffmeyer *et al.* (1986) claimed that the quality increased up to the fourth round. But Reeves and Jauch (1978) argued that opinions hardly change beyond two rounds. There are practical considerations such as higher number of rounds might lead to a high attrition rate (Keeney *et al.*, 2011),

which could weaken the value of the subsequent rounds due to loss of a full range of perspectives.

Rowe and Wright (2001) raised an empirical question that the accuracy improvements seemingly achieved in later rounds may simply be due to the busiest or most impatient panellists dropping out. Additionally, research does show that when responses show reduced variability, this should be accepted as a criterion of agreement, and it is up to the facilitator to decide when to call the procedure to a halt (Rowe and Wright, 2001). In this study, two rounds of Delphi were sufficient to achieve quality responses, reducing the risk of high attrition rates.

The debate about Delphi's epistemological stance is often neglected with no real agreement in sight, but constructivism is often mentioned (*Keeney et al., 2011*). However, Delphi resonates with the participatory principles of critical pragmatism philosophy by fostering change through cooperative inquiry where different groups can learn how to contribute to their democratic society. Dewey (1966) explained that it is necessary to experience another person's opinions and beliefs in order to more fully develop one's own. This participatory principle in critical pragmatism provides a close fit to Delphi since it is a process of individual feedback about a group opinion with opportunities for respondents to change their position primarily on the basis of that feedback.

4.4.3 Evaluation research

The term 'evaluation' denotes judging the value of a number of activities, and Weiss (1972, p. 1) refers to it as "an elastic word that stretches to cover judgements of many kinds". Evaluation can be formative leading to refinement and modification, or summative to determine impact or outcome (Clarke and Dawson, 1999). Lincoln and Guba (1986) added that the term 'evaluation' has 'value' in its root comprising two distinct aspects of value that can be established: merit (intrinsic, context-free value), and worth (extrinsic, context-determined value). They suggested that, depending on the purpose, summative/formative and merit/worth may be taken as

independent to one another to generate four types of evaluation described in Table 4.3.

Table 4.3: The four dimensions of evaluation (Lincoln and Guba, 1986).	
Formative merit	Performed to modify or improve the intervention under evaluation while it is in process of development.
Summative merit	Performed to certify or warrant its merit against a set of standards after the intervention has been developed.
Formative worth	Performed to facilitate adoption or fitting of the evaluation to some local context of use.
Summative worth	Performed to determine the impact of an intervention to warrant or certify it for permanent use.

Based on this classification, Lincoln and Guba (1986, p. 550) defined evaluation as: “A type of disciplined inquiry undertaken to determine the value (merit/worth) of some entity...in order to improve or refine a programme or a system to assess its impact.”

This definition asserts that evaluation is undertaken to establish value and, like research, is a disciplined inquiry. Lincoln and Guba (1986) used the hybrid term ‘evaluation research’ for studies that apply scientific procedures and publicly confirmable logical processes to the collection and analysis of information about the content, structure and outcomes of a planned intervention to make the conclusion credible.

Evaluating the formative worth (utility) and summative worth (impact) of the scoring rubric fitted the objectives of this phase of the research in order to determine its overall effectiveness and impact on the assessment process. This encompasses evaluating the utility of the consensus-based scoring rubric in terms of its usability, clarity and practicality when mentors and students use it, as well as its impact on providing an effective assessment tool to assess competencies in the practice setting. The usefulness of the scoring rubric in enhancing learning, feedback provision, and self-assessment were also evaluated in line with the principles of assessment *for* learning and authentic assessment. Additionally, the evaluation research method fits well with the critical pragmatism principle that

alternative ideas should be “tested by acting upon them ... to guide and organise further observations, recollection and experiments” (Dewey, 1966, p. 160).

In terms of methodology, evaluation research is a form of applied research that aims to produce information about implementation, operation and ultimately effectiveness of a programme designed to bring about change; there is no single research strategy unique to evaluation research or a methodology of its own (Clarke and Dawson, 1999). Evaluation research makes use of many research methods in order to answer research questions about need, efficiency, effectiveness, appropriateness and acceptability (Moule *et al.*, 2017). This view echoes Rogers *et al.*'s (2015) claim that there is no one right way to conduct evaluation research and what is needed is a combination of methods and designs that suit the particular situation. They added that when choosing these methods and designs, three issues need to be considered: the available resources and constraints; the nature of what is being evaluated; and the intended use of the evaluation.

Since evaluation research relies on social science research methodologies, information is obtained by making extensive use of a wide range of well-established social research methods (Weiss, 1998). A consensus exists with respect to the fact that both quantitative and qualitative methods have an important place in evaluation research (Clarke and Dawson, 1999).

A goal-orientated approach, that uses experimental methods to measure the extent to which an intervention has achieved specific goals and objectives, is traditionally used in evaluation research (Moule *et al.*, 2017). Although true experiment is the design of choice in many circumstances due to inherent guards against threats to validity (Weiss, 1998), in this study true experiment design was not considered suitable as undertaking and maintaining random allocation was not possible due to the need to pair up students and mentors. Moreover, it was felt that undertaking a true experiment was not appropriate when comparison between groups was not one of the study's objectives.

Qualitative methods on the other hand are becoming increasingly popular and widely accepted in evaluation research. They offer a vastly increased repertoire of methodological tools that renewed interest in the strategies of integration (Greene and McClintock, 1985). Khandker *et al.* (2009) added that qualitative methods generate information that may be critical for understanding underpinning mechanisms and provide details as to whether the beneficiaries truly benefited from the intervention under evaluation.

A comprehensive evaluation, according to Clarke and Dawson (1999), is not simply identifying quantitative indicators that can be used to measure success but is also about exploring the views, perspectives and the nature of the interaction that takes place. Thus, both quantitative and qualitative research methods should be used when conducting evaluation research, and mixed methods are now an established feature of evaluation research (Clarke and Dawson, 1999). Rogers *et al.* (2005) supported this view, explaining that in any impact evaluation a combination of different methods is needed to answer different types of questions, stressing that impact evaluation is generally most reliable and valid when it uses a mixed methods approach where results from one method can be used to test or extend those of another.

Applying this to social research, Denzin (1970) explained that there are two types of methodological integration: the 'between methods' approach, referring to the actual mixing of methods, and the 'within method' approach, entailing the use of multiple techniques within a given method. The latter was adopted in this study through administering questionnaires using a mixture of attitudinal measures to collect quantitative data and open-ended questions to collect qualitative data (see section 6.2).

4.5 Chapter Summary

This chapter has critically examined the philosophy underpinning the decision that critical pragmatism was the most suitable ontological and epistemological approach as it fosters cooperative and social inquiry to find a practical solution in deciding

what is *truth* and *knowledge* (Dewey, 1966). This philosophy facilitated a pragmatic strategy for the methodological approach of mixing multiple research methods needed to answer different research questions. It also ensured clear alignment between the research questions, design and the pragmatists' principles that methods are selected for their practical consequences and empirical findings without the constraints of following a particular tradition (Brannen, 2005).

There were two phases of data collection that occurred sequentially (see Figure 4.2). The first phase employed Delphi as the most appropriate consensus method using exploratory sequential mixed methods design. Phase two employed an evaluation research method using convergent mixed methods design.

This research is unique at several levels. This is the first study of practice-based assessment to adopt the philosophy of critical pragmatism as well as using Delphi to design a scoring rubric to interpret professional attitude, behaviour and responsibility statements within nursing. The next two chapters provide the actual design and results for each phase of the study. The Delphi phase is presented in chapter 5. Chapter 6 presents the evaluation research phase.

Chapter 5: The Delphi

5.1 Introduction

The IR in chapter 3 identified that within the context of practice-based assessment in pre-registration nursing education, mentors lacked understanding of competency statements or making decisions regarding the appropriate level of competence.

These findings highlighted the need to establish a transparent and common language to interpret different levels of competence to help mentors define what is expected of students and for students to identify what they are expected to achieve.

Building on the theoretical principles of assessment *for* learning and authentic assessment (discussed in section 3.4), the objective was to adopt a pragmatic approach by designing a scoring rubric intended to reduce confusion and ambiguity in interpreting and measuring levels of competence. Thus, help mentors not only to make valid judgements and provide constructive feedback, but also to promote students' self-assessment.

This chapter describes the method used in phase one of the study to answer the research question relating to designing the scoring rubric. In line with the critical pragmatism tenets, the research question had a practical problem-solving approach that values social action and participation to serve the public good as a tool to change (Kadlec, 2007). Here, Delphi is utilised as a social democratic procedure that emphasises social involvement in interpreting different levels of competence and the development of a scoring rubric. The chapter presents the two rounds of Delphi conducted to interpret the professional attitude, behaviour and responsibility statements within the PLPAD for third year pre-registration student nurses. Since the findings from round one of Delphi were used to design the questionnaire in the second round, and the results of the second round informed the development of the scoring rubric, the results for both rounds are presented in this chapter.

5.2 Study design

The study adopted an exploratory sequential mixed methods design consisting of a two-round 'classic e-Delphi' presented in Figure 5.1. Justification for selecting sequential mixed methods and Delphi method as the most suitable approaches in this study was presented in sections 4.3.1 and 4.4.2 respectively.

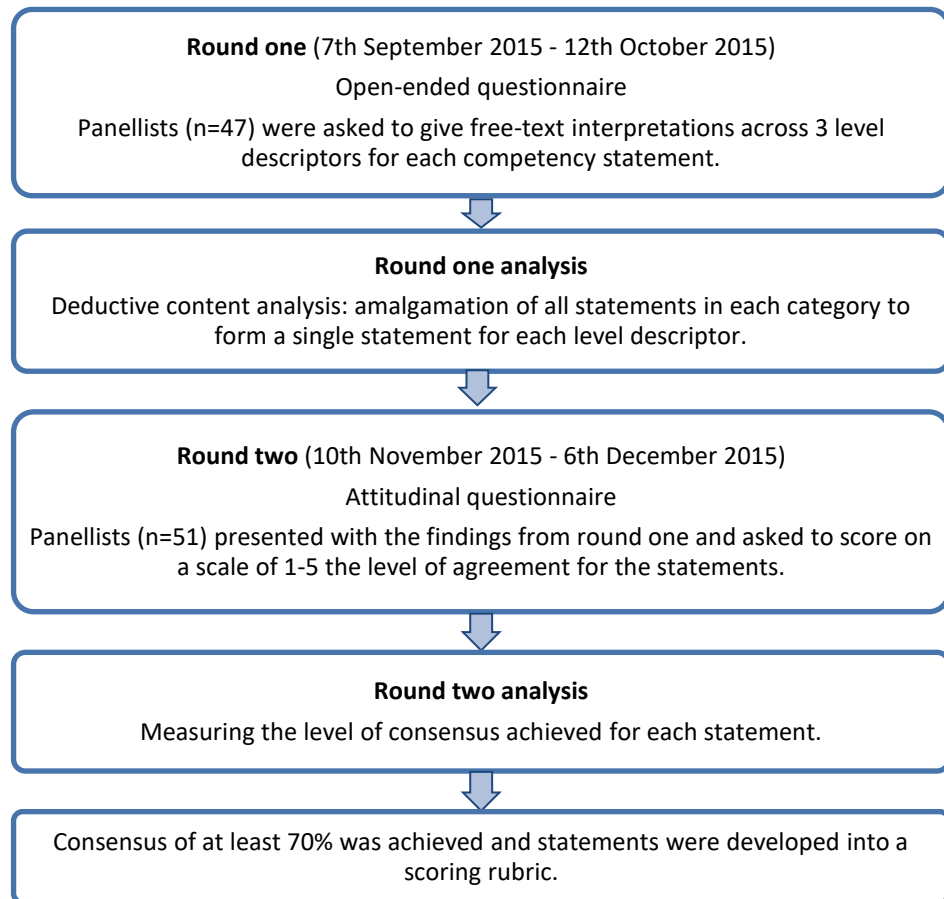


Figure 5.1: Overview of the e-Delphi rounds.

Considering no validated questionnaires found in the literature that could meet the objectives of this study, questionnaires were developed in each of the two rounds of Delphi which used Bristol Online Survey (BOS) website as a host. The first round gathered stakeholders' interpretations of the competency statements, where a qualitative approach using free-text questions was adopted. The second round used data from the first round to formulate attitudinal statements to establish the degree of consensus or divergence in the stakeholders' interpretations of the level descriptors.

5.3 Sample/Sampling

Creating the expert panel

Recruiting a heterogeneous sample was central to ensuring a broad range of stakeholders' perspectives were sought. Representativeness, according to Keeney *et al.* (2011), is based on the qualities of the expert panel therefore sampling techniques such as purposeful, convenience and snowballing are appropriate. Bowling (2014) advised that purposive sampling is appropriate where respondents are selected because they have knowledge that is valuable to the research. Therefore, in order to include a diverse but representative sample in this study a purposive sampling approach was selected to recruit a group with particular characteristics.

Although results cannot be generalisable to the wider population since the sampling was not randomised (Bowling, 2014), Patton (2002) argued that criterion sampling, as a purposeful sampling strategy, shares many characteristics with random probability sampling. Despite having different aims and different procedures for identifying and selecting potential participants, in both instances, study participants are drawn from agencies, organisations or systems involved in the implementation process. This argument provided adequate justification that criterion sampling was the most appropriate strategy by selecting all individuals who met specific predetermined criteria defined on the basis of their role and involvement in pre-registration nursing education.

The definition of 'expert' in Delphi methods remains ambiguous and includes: informed individual (Mckenna, 1994), specialist in the field (Goodman, 1987), or knowledgeable about the subject (Lemmer, 1998). The most fitting definition of the term 'expert' was suggested by Keeney *et al.* (2011, p. 24) as "anyone with relevant input". Guided by this definition, the criteria for inclusion in this study entailed all participants who have knowledge and experience in assessing the practice-based component of pre-registration education, being assessed or having a vested interest in the process, as well as having sufficient time and willingness to participate. Five groups of stakeholders were identified: practice-based assessors, nursing students,

nurse educators, service users and strategy or policy shapers. Pre-determined selection criteria specific to each group is summarised in Table 5.1.

Table 5.1: Selection criteria specific to each stakeholder group.			
Stakeholder	Rationale for inclusion	Inclusion criteria	Recruited from
Nursing students	Assessed through/have experienced repeated assessments using the PLPAD.	Third-year adult field pre-registration nursing students at the HEI.	Student nurses studying at the HEI.
Practice-based assessors (mentors, sign-off mentors or practice educators)	They undertake the role of assessing students in practice placements using the PLPAD.	Mentors/sign-off mentors who hold a recognised qualification to assess nursing students. Practice educators and Practice Development Nurses.	Nurses attending CPPD modules at the HEI. The HEI mentors' website. Live mentors' register Mentor attending annual updates.
Nurse educators	Involved in designing and moderating the completed PLPAD.	Nurse educators involved in the pre-registration curriculum that uses the PLPAD.	Pre-registration educators from the HEI. Pan London Practice Learning Group.
Service users	Involved in teaching, developing and reviewing nursing curriculum.	Service users registered on the HEI 'Peoples' Academy' database.	Established 'Peoples' Academy' collaboration within the HEI.
Strategic or policy shaper	Involved in various stages of developing the PLPAD including consultation, designing, reviewing, and approving the assessment document.	NMC: professional body responsible for setting assessment standards and professional values competencies.	NMC mailing list.
		Pan London Practice Learning Group: developed the PLPAD used in this study.	Pan London Practice Learning Group mailing list.
		RCN: involvement in producing students/mentors practice guidelines.	RCN mailing list.
		NHS: Chief Nursing Officers, Directors of health education. DH Nurse Directors.	NHS and DH mailing list.
		Authors of relevant literature. Editors of relevant Journals.	Contact details within the Journals.

Although there are many similarities in the competency statements across all nursing fields, there are minor variations in some of the statements to make them

applicable to each specific group and different academic year. Therefore, to avoid sending several versions to the panel, it was decided to choose the professional attitude, behaviour and responsibility statements for the adult nursing field only as they comprise the largest number of students. Selecting third year students was based on the assumption that being in their final year before becoming registered nurses, they would have more exposure and experience of using the PLPAD.

Mentors were included if they held the NMC-approved mentorship course or equivalent that allowed them to join the mentors' 'live register'. This included mentors and sign-off mentors (who make judgements about whether a student has achieved the overall standards of competence required for entry to the register at the end of the programme). Practice educators were also included if they use the PLPAD and their role included supporting students and mentors in the clinical settings.

It is acknowledged that the mentors in the sample would potentially include individuals who failed to fail students and it could be argued that they may produce interpretations incongruent to best practice. However, mentors are key stakeholders who have a legitimate role in this process and not including them could not be justified. This potential limitation was overcome by the iterative and controlled feedback characteristic of the Delphi, allowing participants to refine or reconsider their opinions from round to round based on the opinions and judgements of all group members. The collective agreement among all stakeholders to reach a consensus minimised the risk of poorly interpreted competencies.

Nurse educators were included if they were involved in the use or development of the PLPAD. This made educators from across London eligible to participate. However, for recruiting outside the HEI, it was decided to approach educators through the Pan London Practice Learning Group as it has representatives from all universities in the London region who used the PLPAD. Logistically, this provided access to educators from several HEIs in and around London without having to seek ethical approval from their organisations individually.

The service users included in the study were recruited from an established citizens' collaboration within the HEI called 'Peoples' Academy', giving people receiving care or who support them, a voice. Their lived experiences are recognised in transforming education by involving them in recruiting, teaching and assessing students as well as the impact they have on advising and transforming the health and social care sector. Involving service users in reaching a consensus on how to interpret the competency statements not only ensures that the language used is clear and explicit, but that it reflects how a service user perceives competent practitioners. This provides a close link to the assessments *for* learning and authentic assessments theories underpinning this research.

Recruiting citizens from the Peoples' Academy who are involved in educating students and qualified nurses provided strength to the study as they were likely to have more insight and interest in nurse education as well as having vested interest in ensuring unsafe practitioners did not become registered practitioners. From a practical point, accessing service users through Peoples' Academy did not require further ethical approval from other organisations such as the NHS.

Sample size

The size of expert panels is also open to debate as the literature provides a wide spectrum ranging from 4 to 3000 (Cantrill *et al.*, 1996), but most published studies reported a sample size between 10 and 100 (Atkins *et al.*, 2005). On the other hand, Powell (2003) stated that the quality of participants is more important than the size and Keeney *et al.* (2011) suggested 10-15 participants if the sample is homogenous advising that for heterogeneous samples more subjects are required but they did not recommend a figure.

A consideration that needed to be factored in when deciding the sample size in Delphi was the high attrition rate, which is frequently associated with Delphi in the literature (Evans, 1997; Donohoe and Needham, 2008). Primarily this is caused by fatigue due to the commitment required for several rounds of questionnaires (Keeney *et al.*, 2011), distraction between rounds or disillusionment with the

process (Donohoe and Needham, 2008), or finding the exercise more burdensome than anticipated (Hill and Fowles, 1975). Evans (1997) claimed that attrition is mainly associated with large Delphi panels, but made no attempts to define what constitutes a large panel.

A high attrition rate can lead to response bias (Keeney *et al.*, 2011). This may mean that the final results are based on an unrepresentative subset of the original sample and represents a threat of selection bias (Hill and Fowles, 1975). Nonetheless, there was no literature found that looked at non-respondents to analyse reasons for not responding. Reeves and Jauch (1978) discussed how opinions hardly change beyond two rounds, and careful consideration in the planning to avoid repeated rounds may help reduce attrition. Guided by the factors mentioned so far, several strategies were employed in this study to enhance the response rate and reduce attrition and are summarised in Table 5.2.

Table 5.2: Strategies used in the Delphi to reduce attrition.	
Factor	Strategy
Fatigue due to several rounds of questionnaires or burdensome process.	<ul style="list-style-type: none"> • Reducing the number of statements in the questionnaires after the feedback received from the pilot study (presented in section 5.5.2). • Attention was given to the process of analysing data and designing the questionnaires in an attempt to limit the number of rounds (Linstone, 1978). • The questionnaires were piloted to evaluate how demanding they were. • Participants' perception of ownership was enhanced by emphasising the needs for their contribution to design a tool that would enhance practice-based assessment (McKenna, 1994).
Distraction between rounds or disillusionment with the process.	<ul style="list-style-type: none"> • The use of e-Delphi to shorten the time frame between rounds (Donohoe and Needham, 2008). • Detailed information about how to use the questionnaires and what is required were provided on the left side of the screen for every webpage. • The researcher allocated protected time between rounds for rapid collection, analysis and feedback to the panel to ensure speedy process and encourage involvement and motivation (Buck <i>et al.</i>, 1993). • An email was sent to the panel to stay in touch, update them on the data analysis progress and alert them to round two. • The participant information sheet was piloted with a sample similar to the target population to ensure clarity of the instructions.

On deciding the sample size for this study, considerations were given to the importance of recruiting a heterogonous sample to ensure a spectrum of opinions was presented and the size was large enough to mitigate for the risk of attrition

(Donohoe and Needham, 2008). Equally, the data generated needed to be manageable as large panels also make analysis more time consuming with limited benefits (Reeves and Jauch, 1978; Skulmoski *et al.*, 2007). Therefore, 100 potential participants were invited with the aim of recruiting ten participants from each of the five groups shown in Table 5.1.

Due to the quasi-anonymity nature of the Delphi, it is was not possible to identify who took part in each round with the possibility that samples may differ between rounds. This was evident in this study where 47 participants responded in round one and 51 in round two (see section 5.6). While they might be different participants, they all share the same characteristics drawn from the same homogeneous group that was originally invited to participate. Therefore, although there was a slight difference in the number of participants, they were representative of all stakeholder groups identified in Table 5.1, hence very unlikely to have an impact on the overall consensus achieved.

5.4 Instruments

There are several web-based survey tools available on the internet that could have been used to conduct the e-Delphi. However, due to the unique and specific need for questionnaires that list the eight professional attitude, behaviour and responsibility competency statements in the PLPAD, finding an existing questionnaire to meet the study needs was considered very unlikely. Therefore, an online questionnaire was specifically designed for each Delphi round using BOS website as a platform. The decision to use the BOS as the host to design and administer the questionnaires was mainly influenced by the assurance that the website complies with all UK data protection laws and the survey was password-protected with only invited participants able to access the questionnaires. The BOS had the advantages of requiring minimal training and the facility to develop a variety of questionnaire formats including open ended, closed and Likert-style questions that met the needs of all the research phases. Additionally, descriptive statistical analysis is also possible through BOS as well as the facility to export data to other statistical software.

5.4.1 Round one questionnaire design (qualitative)

Round one adopted the classic e-Delphi approach (discussed in section 4.4.2) comprising a qualitative open-ended round to ensure the panel members were provided a space to express themselves in their own words (Murray and Hammons, 1995). The approach ensured participants were given complete freedom in their responses and thereby identify comprehensive possibilities (Hasson *et al.*, 2000). This was thought to be useful for exposing individual differences in interpreting the competency statements among the stakeholders and offered the opportunity to condense collective subjective perspectives on how the statements should be interpreted across three level descriptors (Keeny *et al.*, 2011).

Sackman (1975) explained that careful consideration must be given to avoid vague questions which can result in ambiguous responses. Although this is plausible, the statements provided to the panel were copied from the PLPAD verbatim, and any ambiguity in understanding the statements would reflect the findings from the IR, further indicating the need for this research study. The questionnaire originally included all fourteen professional values statements in the third year PLPAD, presented in Appendix 1. After the pilot, this was revised down to the eight statements specific to the professional attitude, behaviour and responsibility competencies section, as well as dividing the panel into two groups with each group asked to interpret four statements only (see section 5.5.2). To enable the identification of views and interpretations without any influences, three free-text boxes that indicate the different levels of performance were allocated to each statement. In the literature the number of level descriptors typically ranges between 3-5, and the justification for opting to use 3 levels was based on the argument that the more levels there are, the more difficult it becomes to differentiate between levels, hence, fewer levels enhances assessors' reliability and efficiency in discriminating between performance levels (Wolf and Stevens, 2007).

To benchmark each performance level, the 'Successful Outcomes Markers' model comprising the headings 'Don't want to see', 'Expect to see' and 'Love to see' was used, on the premise that this model offers a new approach that focuses on the

hard-to-measure human behaviours (Burkhart-Kriesel *et al.*, 2011). The panel were given the opportunity to comment on the suitability of the Successful Outcomes Markers model in round two (see next section).

The BOS webpage started with a welcome message screen followed by three screens to collect demographic data. Then on each of the remaining screens, at the top of the screen, a message stated: “In your interpretation of the statement below, describe what the expected performance looks like making distinction across the 3 levels”. Below that, alongside each competency statement, they were presented with 3 free-text boxes with the headings: ‘Don't want to see’, ‘Expect to see’ and ‘Love to see’. Further clarification for what each level meant was provided when the participant clicked on the ‘more information’ tab. At the end of the questionnaire, there was a screen thanking participants and reminding them that there would be a second round. It was anticipated that the amount of time for completing the questionnaires would vary between panellists and probably take 30-45 minutes (see Appendix 5 for the round one questionnaire).

5.4.2 Round two questionnaire design (quantitative)

The questionnaire for round two aimed at gauging the stakeholders’ agreement to the analysed level descriptors developed in round one. There are different styles that can be used when developing a questionnaire to produce different types of data. Thus, it is important to be clear which scale and response format to use as this would influence the analysis options (Rattray and Jones, 2007).

The type of data in round two were ordinal to enable the variables to be ranked on a scale of increasing magnitude. Attitudinal Likert scales ranging from strongly agree to strongly disagree are commonly used to collect this type of data (Wood *et al.*, 2011). To mitigate for the restrictions of fixed choice responses and to encourage the panel to share their opinions and suggestions, a free-text box was added to allow for any comments they wish to add.

In using Likert scales, some controversy exists as to whether a neutral point should be offered. Bishop (1987) reported that a neutral category should be included when there is a possibility that respondents may legitimately not know the answer. Gray *et al.* (2017) also noted that removing this option forces the respondent to choose another response, which may lead to respondent irritation and may increase non-response bias. Considering the debate around the neutral option, the decision was made to include this option in the questionnaire to allow those who genuinely had no opinion to indicate this. It was also decided that the neutral category would be placed in the middle rather than at the end of the scale as Bishop (1987) found that respondents were likely to choose the neutral option if it was placed at the end.

Similar to round one, the BOS website started with a welcome message screen followed by three screens to collect demographics data. Every participant was then presented with all eight statements, each in turn being presented in a screen where this time round, the statements were provided at the top of the webpage and below it the three level attributes were provided. Across each level, a five-point Likert scale was positioned ranging from strongly agree to strongly disagree. The questionnaire can be viewed in Appendix 6.

As discussed in the previous section, the draft scoring rubric adopted the Successful Outcome Markers approach 'Don't want to see', 'Expect to see' and 'Love to see', and the panel members were given the opportunity to rate their preferred benchmarking style among the commonly used level descriptor methods presented in Figure 5.2 overleaf. The final screen thanked the participants and alerted them to the possibility of a third round. It was anticipated that it would take 15-20 minutes to complete the questionnaire.

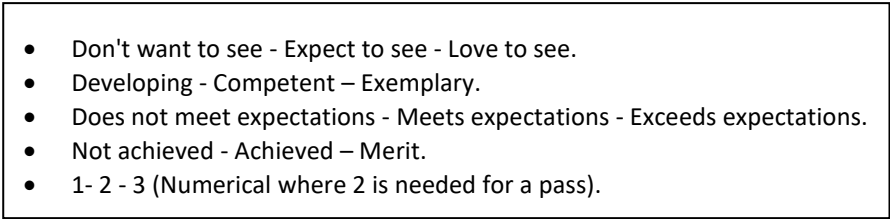
- 
- Don't want to see - Expect to see - Love to see.
 - Developing - Competent – Exemplary.
 - Does not meet expectations - Meets expectations - Exceeds expectations.
 - Not achieved - Achieved – Merit.
 - 1- 2 - 3 (Numerical where 2 is needed for a pass).

Figure 5.2: Commonly used bench markers for level descriptors.

5.4.3 Methodological rigour

As with all types of inquiry, consideration needed to be given to the reliability and validity of the Delphi to evaluate its rigour and integrity. Reliability refers to the degree to which the research method produces stable and consistent results and is approached principally through the standardisation of research procedures (Gray *et al.*, 2017). Problems with reliability of the Delphi method are commonly reported in the literature and predominantly relate to the variations of the procedural applications of the method, which makes standardisation of the method to evaluate its reliability difficult (Woudenberg, 1991).

In the light of potential reliability shortcomings, Hill and Fowles (1975) suggested a set of considerations to enhance methodological rigour in Delphi. Table 5.3 summarises these considerations including how they were addressed in this study.

Table 5.3: Reliability considerations (Hill and Fowles, 1975) and how they were addressed in this study.		
Delphi aspect	Reliability consideration	Addressed in this study
Clarity of questions	Questions must be carefully designed to provide unambiguous stimuli to the respondents and avoid bias or distort responses.	The questionnaires for both rounds were piloted prior to the launch to ensure clarity of the questions.
Choice of respondents	Selection bias by relying on participants who are readily available or professional associates.	Efforts were made to recruit a representative sample typical of the general population of stakeholders.
Character of round one	Depriving the panel of the intended role of expressing their own views and provide them with preselected statements.	Round one followed the classic format and provided the panel with free-text to provide their own views.
Administration of the questionnaires	The use of mail questionnaires is associated with low response rate and if lengthy, respondents find it burdensome.	The questionnaires were administered online allowing the process to be faster to keep respondents engaged in an attempt to improve response rate. Round one questionnaire was made shorter after the pilot test.
Consensus	Ambiguity about level of consensus and the criteria for evaluating the extent of consensus	Consensus in this study was clearly stated in advance as achieving at least 70% agreement on the interpretations of the statements.

Validity relates to how well an instrument measures what it is intended to measure. Hill and Fowles (1975) claimed that validity is enhanced in Delphi given that the method is based on a group opinion, and the decisions are then strengthened by the reiteration and controlled feedback process to challenge assumptions. There is strong evidence in the literature that Delphi provides content and face validity

based on confirmative judgements of a panel that have knowledge and interest in the topic (Keeney *et al.*, 2011).

Criterion validity with its two types, concurrent and predictive, is when a test is shown to be effective in predicting indicators of a construct. Hasson *et al.* (2000) argued that concurrent validity is achieved through reaching a consensus using successive rounds of the questionnaires, and predictive validity is measured in terms of accuracy of the Delphi (Von der Gracht, 2008), and is proof of the method validity (Keeney *et al.*, 2011). By recruiting a sample of stakeholders that is representative of the target population and subsequent evaluation of the developed scoring rubric in different NHS Trusts (see section 6), internal and external validity can be enhanced. However, validity is ultimately affected by response rates (Hasson *et al.*, 2000).

The debate over the methodological rigour according to Day and Bobvera (2005) is ongoing and they questioned the appropriateness of using a positivist approach to measure rigour. Holloway and Wheeler (1996) and Day and Bobvera (2005) believe that Delphi overlaps both quantitative and qualitative approaches, so advocate using the term 'trustworthiness' to determine the appropriateness and effectiveness of a Delphi. Trustworthiness, according to Lincoln and Guba (1985), comprises credibility, dependability, confirmability and transferability. As this study adopted a mixed methods approach, efforts were made to achieve trustworthiness. Table 5.4 (overleaf) summarises these considerations including how they were addressed in this study.

Table 5.4: Trustworthiness considerations (Lincoln and Guba,1985) and how they were addressed in this study.		
Trustworthiness aspect	Consideration	Addressed in this study
Credibility	Confidence in the 'truth' of the findings.	<ul style="list-style-type: none"> • Adopting a well organised research method. • Member verification through iteration and controlled feedback. • Representative sample of stakeholders involved in nurse education. • Debriefing sessions between researcher and supervisors.
Dependability	Showing that the findings are consistent and could be repeated.	<ul style="list-style-type: none"> • Employing quantitative and qualitative methods. • Member verification. • Detailed reporting of methodology to allow the study to be repeated.
Confirmability	Show that the findings are shaped by the respondents and not researcher bias.	<ul style="list-style-type: none"> • The researcher maintained the 'outsider observer' view. • Triangulation of quantitative and qualitative methods. • In-depth methodological description. • Recognising limitations in the study and trying to minimise their potential effects. • Maintaining an audit trail.
Transferability	Demonstrating that the findings have applicability in other contexts.	<ul style="list-style-type: none"> • Providing contextual background and detailed description of the study questions to allow comparisons to be made. • Piloting the scoring rubric to validate the applicability of the Delphi findings (chapter 6).

5.5 Procedure

This section describes how the two rounds of e-Delphi were conducted, including piloting, data collection and data analysis. The two rounds took place from 7th September 2015 to 6th December 2015.

5.5.1 Ethical considerations

Prior to starting data collection, ethical approval for the study was granted from the University Research Ethics Committee (Appendix 7). When the questionnaire for round two was developed based on the data analysed from the first round, it was submitted as a supplement to the original ethics application and was approved as an amendment to the original ethics application (Appendix 8). The PLPAD is copyrighted, therefore permission from the Chair of the Pan London Practice Learning Group to feature the professional values statements in the thesis as well as providing a copy of the competency statements in the appendix was obtained.

Participants were adults with sufficient capacity to make their own decisions and potential harm in this study was considered unlikely due to the type of data they

were required to provide. However, they were reminded that participation was voluntary and that they were free to withdraw at any point without the need for explanation. They were also informed that their data would no longer be retrievable once it had been anonymised. In the unlikely event that participants may have experienced any distress, they were reminded they could stop completing the questionnaire and encouraged to contact the researcher. Service users were recruited from the Peoples' Academy within the HEI and could have been referred to staff counselling services.

The main ethical considerations in this study related to confidentiality and anonymity. This was mitigated for through using a secured website for data collection where only invited participants allowed access. Responses were downloaded to a computer to allow analysis, which was encrypted and password protected. Participant identifying numbers were used, only known to the PhD supervisory team.

Anonymity is beneficial for group members who may be anxious to share their views in a group as they will be inclined to give more honest answers (Linstone and Turoff, 1975; Murphy *et al.*, 1998). Anonymity gives panellists the opportunity to change their opinions without fear of "losing face" in the eyes of the others in the group (Rowe and Wright, 2001, p. 126). In contrast, Sackman (1975) suggested that anonymity may result in a lack of accountability of views expressed and may encourage snap decisions. It was felt that the sequential and iterative nature of the Delphi and the professional background of many of the participants would discourage this.

True anonymity was not possible in this study since some participants were known as they had been purposefully selected (Keeney *et al.*, 2011). However, 'quasi-anonymity' where respondents may be known but their opinions remain anonymous (McKenna, 1994) was achieved and direct quotes to free-text answers used as part of the study report or later Delphi iterations were not traceable. Although participants' identity was not disclosed even to each other, there is a

chance that participants may have talked among themselves about the process and results. This would have been their autonomous right.

5.5.2 Piloting the online questionnaires

Pilot testing questionnaires is an important element to identify ambiguities in the wording, readability and content (Turoff, 2006), and to identify and adjust any technical issues with accessing the website, successful completion and submission of the questionnaire by members within and outside the HEI (Bowling, 2014). The questionnaires were piloted with a sample (n=8) representative of all five groups presented in Table 5.1. Since the aim of the pilot was to check practical and technical issues such as clarity, time it took to complete the questionnaire and ease of access to the website rather than providing interpretations of the competency statements, the participants in the pilot were excluded from the main study.

Two issues arose from piloting the questionnaire for round one. The first issue was related to a service user who did not understand what the word 'rubric' meant. This was a minor issue and the participant information sheet was modified to include a definition. The second issue, reported by most participants, was related to the length of the questionnaire. They found that interpreting three levels for all fourteen professional values competency statements in the PLPAD to be very lengthy and labour intensive. This raised concerns over the potential risk of high attrition. This was a critical finding and may have had a significant impact on the study. In consultation with the supervisors and those who participated in the pilot, the decision was made to reduce the number of competencies from fourteen to eight.

The decision to choose eight statements was because the professional values statements in the PLPAD are divided into two sections. The first eight statements were under the heading 'professional attitude, behaviour and responsibility', and the remaining statements were under the heading 'safe and compassionate care'. Deciding to choose the professional attitude, behaviour and responsibility statements was influenced by literature reviewed in chapter one, identifying attitudes and behaviours to be notoriously challenging to define or measure

(Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016), thus, the need to have a consensus on how to interpret them would be a practical solution to serve the public good (Dewey, 1966).

To further reduce the time it took to complete the questionnaires, participants were separated into two groups for the first round and each participant was asked to interpret only four competency statements. It is important to note that they were divided equally according to their characteristics to preserve their representativeness. The two groups were merged in the second round so every participant had a chance to review the interpretations for all eight competency statements.

This modification had an impact on the subsequent phase of the study. The original aim was to develop a scoring rubric for all the fourteen statements in phase one, with the intention to evaluate the usefulness of the scoring rubric in the second phase by comparing a group using the scoring rubric to another group using the standard methods currently used. This had to be modified for the pragmatic reasons stated and is discussed in section 6.2.

The questionnaire for round two was also piloted to identify and adjust any readability, ambiguity, or website access issues. The same sample that reviewed the questionnaire in round one participated in the pilot (n=8). Feedback from the pilot indicated that the online questionnaire was easy to access and the instructions and the wording were appropriate with no comments to suggest any changes.

5.5.3 Inviting and retaining participants

Potential participants were invited to take part by email (Appendix 9). The participant information sheet (Appendix 10) was attached to the email which included a brief background, purpose of the study, what was expected of them, how the information collected would be used and reassurance about anonymity and confidentiality of their involvement. These steps help participants feel invested in

the process, resulting in a lower rate of attrition (Hasson *et al.*, 2000). The email also had a link directing them to the BOS website.

The website main page had a welcome message thanking them for taking part as well as the contact details of the researcher should they need assistance. To reiterate informed consent, the following sentence was included “Clicking on the link below indicates that you have read the participant information sheet and you voluntarily agree to participate in the survey”.

The questionnaire for round one was made available from 7th September 2015 to 12th October 2015. Due to the anonymity of responses, reminder emails were sent to all participants. A reminder email was sent after 2 weeks (Appendix 11) and a final reminder a week later (Appendix 12). In an attempt to enhance the response rate, a copy of the participant information sheet was also attached to the reminder emails to provide quick access to the information instead of searching through old emails.

Round two used a summary of the interpretations based on the information provided in the first round. Accordingly, the panel were asked to rate their level of agreement with the statements using a 5-point Likert scale ranging from strongly agree to strongly disagree. Due to the quasi-anonymous nature of the Delphi, it was not possible to identify who responded in round one, therefore, invitation emails were sent to all potential participants whether or not they took part in round one (Appendix 13). The questionnaire was made available from 10th November 2015 to the 6th December 2015. Again, a reminder email was sent after 1 week (Appendix 14) and a final reminder a week later (Appendix 15).

5.5.4 Data analysis (round one)

Acknowledging that a slow Delphi process could negatively affect participants’ engagement and response rate (Hsu and Sandford, 2007; Keeney *et al.*, 2011), proper planning and time management was necessary. Blocks of time were dedicated to the process of conducting the data analysis, the developing of a new

questionnaire based upon the responses, and distributing the subsequent questionnaire in a timely fashion.

The open-ended nature of the first round meant that qualitative analysis was the most suitable approach (Keeney *et al.*, 2011). Methods of qualitative data analysis vary according to the purpose of the Delphi study. Although both content and thematic analysis have similar phases of familiarity with data to obtain the sense of the whole then coding the data under categories or themes, the literature indicated that qualitative content analysis (QCA) techniques were more appropriate for the initial unstructured questionnaire to identify and group statements generated by the panel (Powell, 2003; Keeney *et al.*, 2011).

Elo and Kyngas (2008) explained that the primary aim in QCA is to describe the phenomenon in a conceptual form and the content analyst views and interprets data to act on their meanings and make sense of what is mediated between people. In contrast, thematic analysis applies minimal description to data sets and interprets various aspects of the research topic involving identifying, analysing and reporting of emerging patterns or themes that extend across the entire data set (Braun and Clarke, 2006).

The rationale for choosing QCA over thematic analysis to interpret the data in round one was that the themes were already provided in the form of competency statements. QCA allows the researcher to pull out key words and statements that have similar meanings then group and collapse them into one statement, while ensuring that the statements provided by the panel remain as true to the wording as possible (Powell, 2003; Keeney *et al.*, 2011). In contrast, interpretation of latent content in thematic analysis to generate themes and patterns is likely to be influenced by the preconceptions, assumptions and 'world view' of the researcher (Elo and Kyngas, 2008), which contradicts the objective of using the Delphi to reach a consensus based on the panel's collective interpretations of the statements.

Before starting analysis, the researcher must decide whether to analyse only the manifest content or the latent content as well. The aim with latent content is to notice silence, sighs, laughter, posture etc. (Gray *et al.*, 2017), which would be impossible to capture or observe in a questionnaire. Guided by the aim in this stage of the Delphi, manifest content analysis was more appropriate in order to stay faithful to what the participants had written. It is important to note that in QCA the researcher is still immersed in the data to look for the merit behind the words. QCA is concerned with meanings, intentions, consequences and context rather than superficially looking at the frequency of what was said alone (Downe-Wamboldt, 1992; Cavanagh, 1997).

Within QCA, there are two fundamental approaches to analysing the data: the inductive approach and the deductive approach (Burnard *et al.*, 2008). In the inductive approach, categories are derived from the data; thus, it is used for analysing data with little or no predetermined theory, structure or framework, and uses the actual data itself to derive the structure of analysis.

Deductive content analysis, on the other hand, is used when the structure of the analysis is operationalised based on previous knowledge and the purpose of the study is theory testing (Burnard *et al.*, 2008). This approach is based on an earlier theory or model where researchers are already aware of probable participant responses and wish to retest existing data in a new context (Elo and Kyngas, 2008). Considering that the overall aim of the analysis in this study was to build on the previous work of the Pan London Practice Learning Group on the PLPAD competency statements rather than generating new ones, the deductive content analysis approach was more fitting than inductive content analysis.

There are computer-assisted packages available to manage qualitative data analysis and Nvivo is a commonly used software. Following attendance at a workshop to learn how to use Nvivo, it was realised that software simply manage, sort and organise rather than 'analyse' data. Therefore, whilst computer programmes can facilitate data analysis, they are merely instruments as good or as bad as the

researcher using them (Burnard *et al.*, 2008), therefore an informed decision was made to manage data analysis manually. This decision was also supported by Webb's (1999) assertion that it is preferable for novice qualitative researchers to use manual methods to gain insight into the intuitive aspects of analysis. Basit (2003) also supported this view arguing that manual data analysis allows the researcher to be closer to the data.

Data analysis strategy

In qualitative content analysis, there are no systematic rules for analysing data. The key feature is that the many words of the text are classified into much smaller content categories (Weber, 1990; Burnard, 1996). Elo and Kyngas (2008) examined several approaches for content analysis and concluded that they all share three main phases: preparing, organising and reporting, which were adopted in this study to guide the analysis process.

Describing the analysis is often one of the most challenging phases in Delphi and researchers often wish for more detailed instructions on how to carry out content analysis (Elo and Kyngas, 2008). This was evident in reporting the analysis of round one in this study as presenting how all the data were analysed within the thesis was not practical and would have been repetitive because the same procedure was followed for each of the eight competency statements. Therefore, it was decided to provide an exemplar explaining the analysis strategy. Nonetheless, every effort has been made to maintain transparency of the steps taken throughout the analysis process.

Preparing the data

The preparation phase started with extracting the stakeholders' responses from the online questionnaire transferred verbatim. A table was developed for each of the eight statements where every participant's comments placed in a row against three columns placed across to represent each of the three level descriptors. The data were not corrected or edited for spelling or grammatical errors at that stage, in order to retain the authenticity of what the participants had written. Only

participants' identifying numbers and the group they belonged to were recorded.

Table 5.5 illustrates how the data was transferred and prepared.

Table 5.5: Example of the how data was transferred, prepared and colour coded for similar meaning			
PLPAD statement 1	Unit of analysis: The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and protection from harm.		
	Don't want to see	Expect to see	Love to see
9300624 Strategic/ policy shaper	talking about patients in public environments. leaving confidential documents, such as patient notes and nursing records, lying around. ignoring unsafe practice by others	precise record keeping and respecting patients right to confidentiality. raising concerns and problems to mentors and qualified staff they are working with.	challenge others on there failings to comply with nmc code. champion the patints rights to harm free care
9458685 Service user	I would not want to hear my confidential information discussed within the hearing of other patients nor would I wish to hear confidential information about other patients on the ward. Similarly I would wish to see written information protected	I would expect to see all my details are kept confidential-including my condition, treatment and medication.	I would love to see the nurse fully aware of data protection issues including both the need to keep data secure but also to know the circumstances when data should be shared in my/my family's interest.
9460832 Academic	Names, Venues, Location, dates	documentation of critical incident	Signed to verify
9403725 Practice assessor	Students not maintaining confidentiality as per national/local guidelines/policies.	Students always maintain confidentiality issues.	All students are well aware and adhere to protect the service users/staff/public.
9460097 Student	Students giving out patient information.	Students to maintain patient confidentiality at all times by not giving out personal information.	Students checking who they are giving information out to.

The next stage was to select the unit of analysis. This can be a word or a theme that is representative of the universe from which it is drawn and not further divided in the course of the analysis (Krippendorff, 2012). Guided by this description, the most fitting unit of analysis in this study was each of the eight PLPAD competency statements. The content analysis process initially began by becoming familiar with the data through reading the material several times and become immersed to obtain a sense of the whole (Tesch, 1990; Burnard, 1991). Table 5.5 illustrates how colour coding was used to distinguish the categories of the statements that have similar meaning.

The next stage was to develop units of meaning within each unit of analysis. Units of meaning comprise statements that have similar meaning (Graneheim and Lundman, 2004). To illustrate, the colour coded content in Table 5.5 were categorised into five units of meaning presented in Table 5.6 below.

Table 5.6: Units of meaning developed from the colour coded content.	
Colour code	Unit of meaning
	Handling of confidential information
	Acting within the Code, policies and guidelines
	Sharing information for safeguarding
	Confidentiality with written information
	Others (no clear context)

Organising the data

After categorisation had been developed, the data were then organised into their units of meaning categories with the corresponding level descriptor. At this stage the focus was on what participants said rather than who said it. Table 5.7 illustrates how the data from Table 5.5 were organised into three level descriptors within their corresponding unit of analysis categories.

Table 5.7: Organising the data according to their units of meaning categories across the level descriptors.		
Category	Level descriptor	Meaning units
Handling of confidential information	Don't want to see	<p>talking about patients in public environments.</p> <p>I would not want to hear my confidential information</p> <p>discussed within the hearing of other patients nor would I wish to hear confidential information about other patients on the ward.</p> <p>Students giving out patient information.</p>
	Expect to see	<p>no careless chat in corridors.</p> <p>Students to maintain patient confidentiality at all times by not giving out personal information.</p>
	Love to see	<p>Students checking who they are giving information out to.</p>
Acting within the Code, policies and guidelines	Don't want to see	<p>Students not maintaining confidentiality as per national/local guidelines/policies.</p>
	Expect to see	<p>precise record keeping and respecting patients right to confidentiality.</p> <p>Students always maintain confidentiality issues.</p>
	Love to see	<p>challenge others on there failings to comply with nmc code.</p> <p>champion the patints rights to harm free care</p>

Table 5.7 cont.: Organising the data according to their units of meaning categories across the level descriptors.		
Sharing information for safeguarding	Don't want to see	ignoring unsafe practice by others
	Expect to see	raising concerns and problems to mentors and qualified staff they are working with.
	Love to see	but also to know the circumstances when data should be shared in my/my family's interest. All students are well aware and adhere to protect the service users/staff/public.
Confidentiality with written information	Don't want to see	leaving confidential documents, such as patient notes and nursing records, lying around. Similarly I would wish to see written information protected
	Expect to see	I would expect to see all my details are kept confidential-including my condition, treatment and medication.
	Love to see	I would love to see the nurse fully aware of data protection issues including both the need to keep data secure
Others (no clear context)	Don't want to see	Names, Venues, Location, dates
	Expect to see	documentation of critical incident
	Love to see	Signed to verify

The statements within each of the meaning units were collapsed and condensed into the three descriptors as close to the text as possible and designed into a scoring rubric format draft (Table 5.8).

Table 5.8: Draft scoring rubric for statement 1.			
PLPAD statement 1	Don't want to see	Expect to see	Love to see
The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and self-harm.	Discloses details about patients, colleagues or clinical area in public including social media. Discusses information with staff not involved in the care or within earshot of others. Does not appreciate limits to confidentiality and need to raise concern for safeguarding purposes. Lacks attention regarding the protection and disposal of written information.	Consistently maintains confidentiality in their day to day practice showing correct processes for discussing, accessing, sharing, storing and disposing of sensitive data in all formats. Understands the principles of public protection and individual safeguarding in situations where confidentiality should be breached. Able to illustrate the appropriate processes to disclose information.	Recognises the boundaries of confidentiality in all forms of communications and is skilful in balancing the sensitivity of maintaining confidentiality with the needs of distressed relatives to seek information. Understands the tensions that occur in practice by continuing to challenge others when they fail to comply with the confidentiality and champions the patient's rights to 'harm free' care.

The content analysis was independently verified by the supervisors to ensure rigour of analysis was maintained. There was a significant amount of typing, spelling and

grammatical errors, which the chief researcher and the supervisors edited before the scoring rubric was presented back to the panel in the second round. Attention was paid to ensure that the editing exclusively focussed on correcting errors in English language while keeping faithful to what the participants meant. One of the supervisors is not a nurse, therefore, staying truthful to the content was enhanced through an 'outsider' review. In addition to the overall objective of measuring the degree of agreement among all participants, sending the questionnaire back to the participants for a second round provided an opportunity for them to verify if the data analysis, the level descriptors and editing of the English presented an accurate account of their interpretations of the competency statements.

Reporting

The same analysis process was carried out for each of the eight PLPAD competency statements. The scoring rubric draft after round one incorporating all eight statements can be found in Appendix 16.

5.5.5 Data analysis (round two)

The data analysis in round two aimed at measuring the level of consensus. Inferential statistics were not considered appropriate because although there were different groups participating in the Delphi, the objective was to ascertain consensus within the sample rather than making comparisons between subgroups. Therefore, calculating the percentage of responses for each statement to indicate levels of agreement was considered sufficient. Calculating the percentage of responses of each statement was already provided through BOS website, hence, there was no need to transfer the responses to SPSS to calculate the percentages.

There is no agreed threshold in the literature to determine a cut-off point for reaching consensus, with suggestions ranging from 51% (Loughlin and Moore, 1979; McKenna, 1994) to 100% (William and Webb, 1994). Based on the objectives and criteria for selection in this study, achieving a 70% level of consensus as recommended by Mitchell (1998) and Keeney (2006) was applied. Although this is not based on any theoretical or methodological standards, it is considered a strong

cut-off point for measuring the level of consensus (Keeney *et al.*, 2011). Therefore, consensus was considered achieved in this study when agree or strongly agree responses scored 70% or higher.

5.6 Results

5.6.1 Round one findings (qualitative)

Out of the 100 invitations, a total of 47 questionnaires were returned in round one (23 from group 1 and 24 from group 2). There are no criteria for acceptable response rates or attrition in Delphi. The literature reports response rates ranging from 8% (Cooney *et al.*, 1995) to 100% (Owens *et al.*, 2008), with some authors recommending a 70% response rate to maintain rigour (Bork, 1993; Sumison, 1998). However, achieving 70% requires considerable effort where the researcher must know the identity of respondents, and non-respondents must be pursued, which presents problems with anonymity (Hasson *et al.*, 2000).

Because the invitation email and the link to the questionnaire was a one stage process, it was not possible to differentiate between those who accepted to be recruited and those who chose not to. Therefore, the accurate response rate of those who accepted to take part was difficult to calculate but it can be safely assumed to be at least 47% if all those who were invited agreed to participate. It is very likely that there were some who did not wish to take part, and if they were identified and excluded, the response rate would be calculated higher than 47%. Nonetheless, there was a good spread of responses representative of all groups as shown in Figure 5.3.

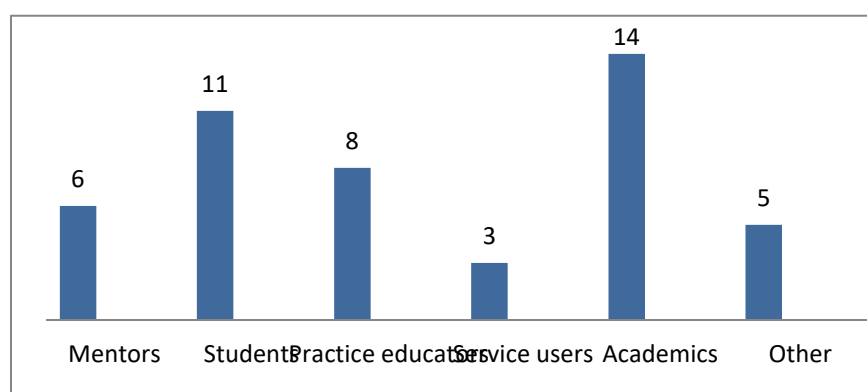


Figure 5.3: Spread of stakeholders' responses in round one

The inclusion of demographic data is not always essential in Delphi studies (Keeney *et al.*, 2011). However, providing a profile of the expert panel in this study gives assurance that the spread of responses shows a good representation of stakeholders. On average 87% of mentors and practice educators had more than five years' mentoring experience and seven identified themselves as a sign-off mentor.

Closely looking at the responses, it was noted that some respondents had answered sub-questions directed at other groups. For example, there were four service users taking part, but sixteen responded to a sub-question aimed at service users only, despite a message on the top of the screen clearly directing them away from the screen if they were not a service user. In another question, a service user identified their discipline as 'other' despite a category already provided for service users. This was adjusted by moving the participant to the service user group. The remaining participants in the 'other' group were a Health Education Dean, a Darzi Fellow, a Lead Nurse for Education, and an NHS Trust Commissioner. One of the participants who selected 'other' entered 'Delphi' in the space provided, making it difficult to identify the discipline. It was also noticeable that none of respondents identified themselves as NMC or RCN affiliate. It would have been extremely useful, as the professional bodies, for them to have shown willingness to engage and have their input.

Overall, the majority of those who took part completed all sections of the questionnaire, but there were a few instances in which participants left sections incomplete. Item non-response is common in questionnaires and occurs when participants fail to provide response to one or more items (Brick and Kalton, 1996). Additionally, six participants (1 service user, 3 students and 2 academics) only completed the demographic pages and did not continue to interpret the competencies. When excluding those six participants, out of a possible 516 statements to be generated only 7% of the data were missing.

There was no answer found in the literature to guide how missing data in Delphi should be handled especially in relation to the qualitative first round. The decision was made that the effect of the missing data on the analysis and findings was negligible, and panellists, including those who partially completed their questionnaires would still have the chance to review and rate the final statements in the second round.

Overall, round one succeeded in its objective to elicit interpretations for the eight PLPAD statements with contribution from the five groups of stakeholders. A total of 24 attributes (three level descriptors per statement) were drafted into a scoring rubric (Appendix 16) ready for round two to seek verification from the panel on their level of agreement to the interpretations.

5.6.2 Round two results (quantitative)

The aim of round two was to measure the stakeholders' agreement to the level descriptors developed in round one. The same Delphi panel members from the first round were asked to rate their level of agreement to the 24 attributes in the scoring rubric using a Likert scale ranging from strongly agree to strongly disagree.

Out of the 100 invitations, a total of 51 participants completed the questionnaire with a response rate of at least 51% (see previous section). There was good representativeness from across all stakeholders' groups as shown in Figure 5.4. It was noted that when comparing the demographics of the two Delphi rounds, differences were evident. On average, 72% of all mentor and practice educators had more than 5 years' experience and eight identified themselves as a sign-off mentor.

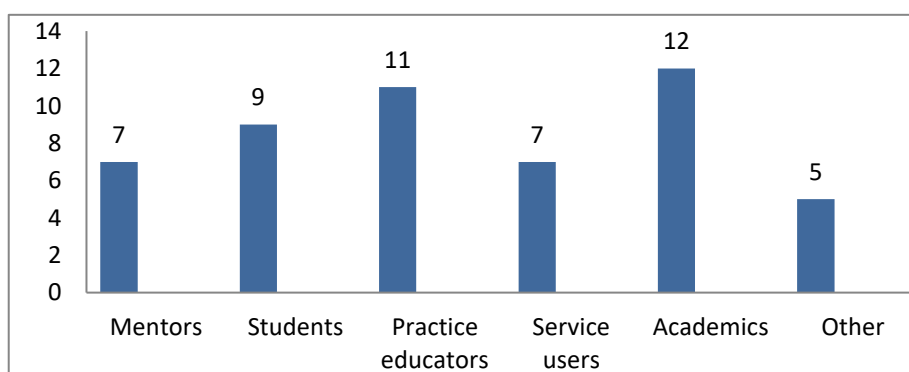


Figure 5.4: Spread of stakeholders' responses in round two

The responses clearly divided opinions as there were very few undecided responses (0-4%). However, there was evidence that respondents experienced difficulty in appraising negatively worded statements in the questionnaire. For example, two respondents consistently misread the questions as evidenced by them strongly agreeing to statements that describe a behaviour considered extremely unprofessional and should have been disagreed with.

There is a great deal of debate in the literature on including or excluding outliers, and what is emerging is that excluding them is subject to the researcher's careful consideration and should be done with caution (Osborne, 2004). However, there seems to be a strong argument not to exclude data if the sample is legitimate (Orr *et al.*, 1991). Outliers were not excluded in this study and kept as part of the data to avoid any concern that data were removed to produce 'desired' results. Their distorting effect on the overall results was negligible.

Similar to round one, there were few inaccuracies in the responses to sub-questions directed at other groups. For example, despite only eight participants identifying themselves as service users, 31 completed the sub-question aimed only at that group. Another eight participants identified their discipline as 'other'; among them there were two service users. The remaining participants in the 'other' group identified themselves as Practice Development Nurse (also ticked 'other'), Health Education Dean, Darzi Fellow, Lead Nurse Education, NHS Trust Commissioner, Senior Nurse Education, joint academic/clinical and education nurse and a Corporate Nurse. Similar to round one, none of respondents associated themselves with the NMC or RCN.

As discussed earlier in (see section 5.5.5), consensus in this study was considered achieved when agree/strongly agree responses scored 70% or higher. All 24 items in the questionnaire achieved a strong consensus with 86%-100% of participants agreeing or strongly agreeing with the statements (consensus results are provided in Figure 5.5 overleaf: the 24 statements can be viewed within the scoring rubric in Appendix 16).

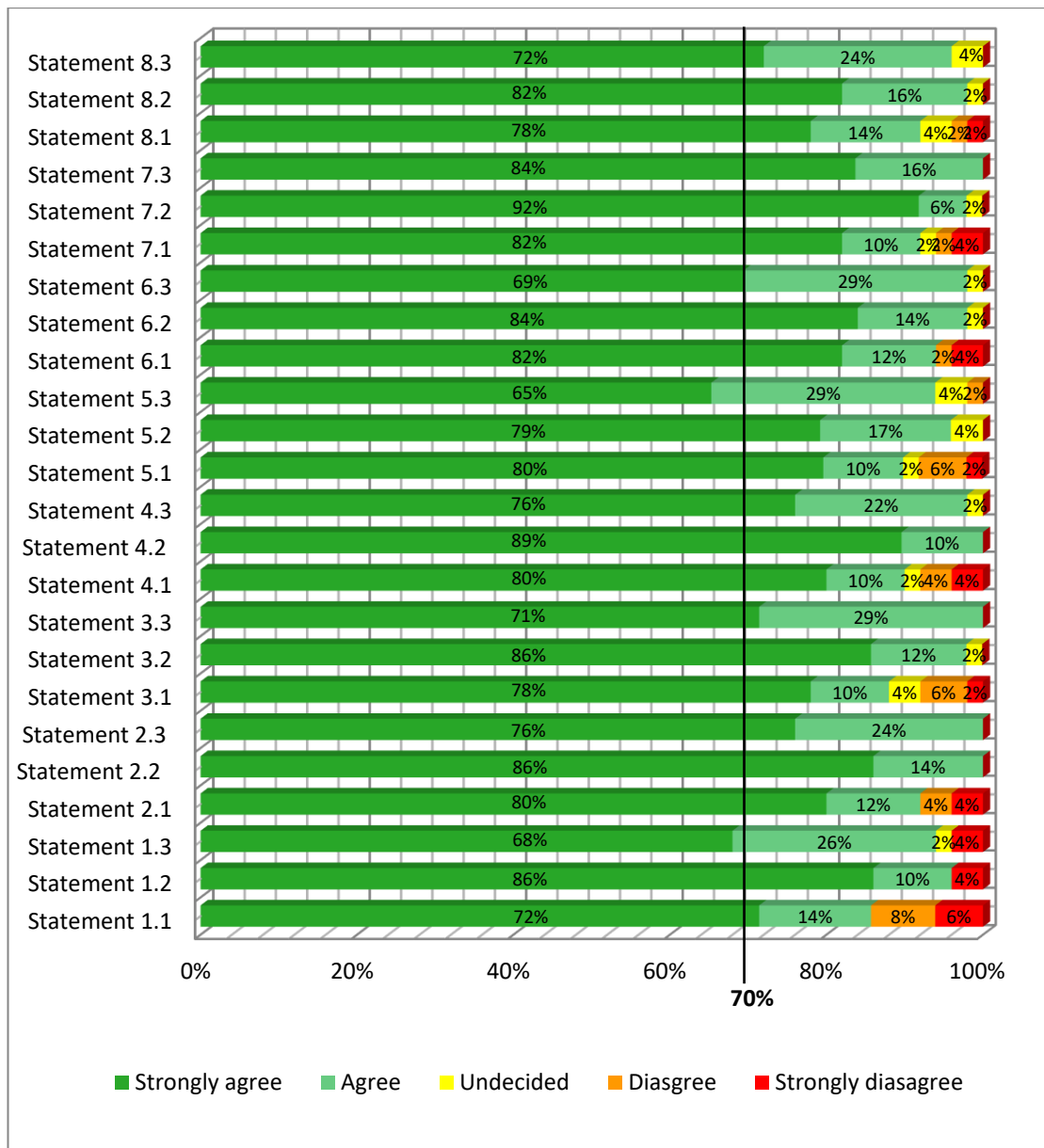


Figure 5.5: Level of consensus for the 24 items in the questionnaire

The strong consensus further supported the decision that descriptive statistics would have not added any extra value. The second round clearly achieved the predetermined consensus level of 70% (ranging from 86% to 100%) in the interpretations of all 24 statements, eliminating the need for a third round.

As discussed previously in section 5.4.2, participants were also asked to rate their preferred terminology to use for benchmarking in the final scoring rubric. They were provided with the commonly used level descriptors. Contrary to the theoretical presumption that the Successful Outcome Markers approach of ‘Don't

want to see', 'Expect to see' and 'Love to see' would be ranked high, most participants preferred to use 'Does not meet expectations' – 'Meets expectations' – 'Exceeds expectations' (Figure 5.6), which was adopted in the final draft.

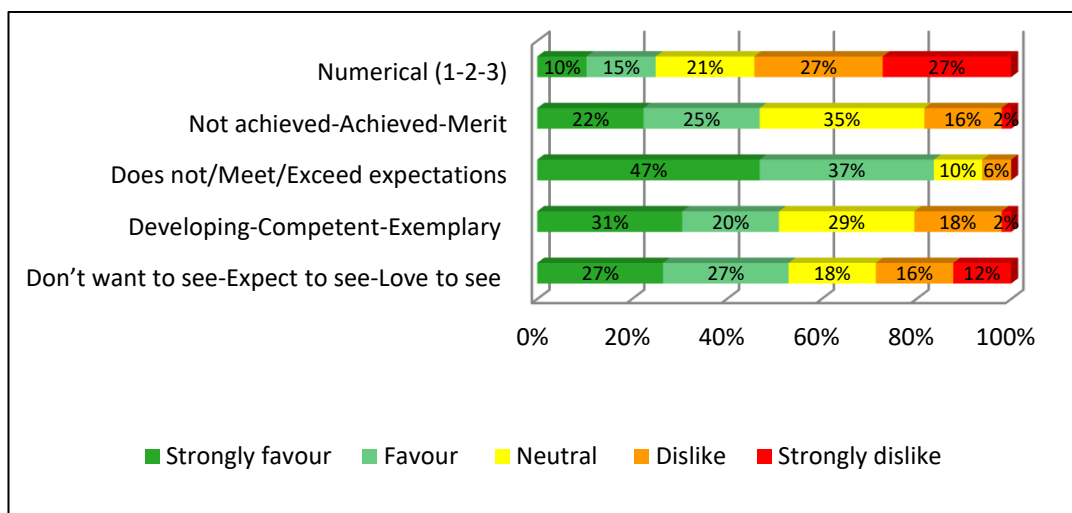


Figure 5.6: Participants' preference for benchmarking terminology

There were no comments provided by the panel members to explain their choices but having the word 'expectations' may have provided the panel with clearer distinction of the levels.

5.6.3 Findings from free-text responses in round two

Participants were given the opportunity to expand on the fixed choice responses using the free-text box provided for making additional comments. These comments provided a unique opportunity for participants to evaluate further the contents of the scoring rubric. In total, 46 comments were recorded, including those related to terminology or clarifying the response provided. Appendix 17 provides a summary of the comments and the actions taken. The comments were examined carefully and taken into consideration and, following the amendments, the revised version of the scoring rubric (Figure 5.7 overleaf), which is the version used in this study, went through a final inspection by the chief researcher and the supervisors before it was administered in the second phase of the research study.

Professional values		Does not meet expectations	Meets expectations	Exceeds expectations
1	The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and self-harm	Discloses details about patients, colleagues or clinical area in public including social media. Discusses information with staff not involved in the care or within earshot of others. Does not appreciate limits to confidentiality and need to raise concern for safeguarding purposes. Lacks attention regarding the protection and disposal of written information.	Consistently maintains confidentiality in their day to day practice showing correct processes for discussing, accessing, sharing, storing and disposing of sensitive data in all formats. Understands the principles of public protection and individual safeguarding in situations where confidentiality should be breached. Able to illustrate the appropriate processes to disclose information.	Recognises the boundaries of confidentiality in all forms of communications and is skilful in balancing the sensitivity of maintaining confidentiality with the needs of distressed relatives to seek information. Understands the tensions that occur in practice by continuing to challenge others when they fail to comply with the confidentiality and champions the patient's rights to 'harm free' care.
2	The student is nonjudgmental, respectful and courteous at all times when interacting with patients/service users and all colleagues.	Judges others based on their differences, e.g. culture, religion, race, gender or their own values and beliefs. Is rude or aggressive to others or ignores their needs. Engages in inappropriate conversations, e.g. gossip, arguments or derogatory comments.	Treats others as individuals and with courtesy irrespective of their diversity while respecting their differences. Unprejudiced, unbiased, and acts professionally in a way that shows courtesy and politeness, both in language and manner.	Actively seeks to understand and embrace the cultural and religious principles of patients and relatives, and advocates upholding such values on their behalf. Recognises own personal perspectives that may influence how they provide care. Is able to reflect on situations that caused them to question their own actions or those of others.
3	The student maintains appropriate professional attitude regarding punctuality and communicates appropriately if unable to attend placement.	Non-attendance or non-adherence to start, finish and break times. Not reporting lateness or non attendance. Changes shifts without good reason and at short notice. Lacks understanding of the importance of punctuality on teamwork and patient care.	Adheres to sickness and absence policies. Consistently punctual throughout the placement allocation. Communicates all attendance issues to placement and university in a professional and timely manner and is proactive in planning and communicating their return following a period of absence. Makes rostering requests in a timely and considerate manner to facilitate any required planned absence.	Illustrates positive attitude and willingness to gain as much as possible from the placement, (e.g. coming prepared for all shifts, planning learning opportunities). Acts as a role model in promoting good work ethic and is conscious of the impact of lateness/absence on patient care, the team they work with and the completion of the course they are studying. Proactive in preparing for the allocated duty and is considerate of the placement needs before their own convenience.
4	The student's personal presentation and dress code is in accordance with the organisation's uniform policy.	The uniform is dirty/creased or incorrect for the organisation. Name badge/ID not displayed or wears uniform outside placement without a reason. Unkempt or unprofessional appearance or poor hygiene. Does not adhere to the organisation's uniform policy or promote a professional image	Consistently looks presentable and clean. Wears the correct uniform for each shift which is clean, well pressed and displaying ID clearly. Able to explain the importance of presentation and dress code and cleanliness and it's association to professionalism and patient safety.	Presents themselves in a professional manner that demonstrates pride in wearing the uniform. Acts as a role model to promote and encourage others to appreciate the importance of professional appearance. Challenges others when their personal presentation or hygiene standards may affect patients' safety or confidence.

Figure 5.7: The final draft scoring rubric

5	The student acts as a role model in promoting a professional image.	Exhibits behaviour, appearance and/or body language that reflects lack of interest in nursing/healthcare. Lack of understanding of how inappropriate behaviour and/or action can impact on the profession and/or the organisation	Sees self as representative of the profession and organisation. Takes pride and sets an example of acting in a way to uphold and promote professional values and passion for nursing in all interactions.	Challenges others who may discredit the profession/organisation. Contributes to enhance the public perceptions of nursing and speaks up when nursing is inaccurately portrayed. Engages in activities and advocates a positive image of nursing as a profession.
6	The student is proactive in promoting and maintaining the individual's privacy and dignity.	Unnecessarily exposes patients during care (e.g. curtains, doors, windows) or disregards privacy measures (e.g. enters or allows others entry without permission). Not taking steps to enable patients (or carers) to carry out own personal needs in privacy.	Demonstrates knowledge and shows consistency in maintaining dignity and privacy in their practice and always takes steps to ensure all measures are taken without being prompted. Fulfils their role in promoting privacy and dignity by correcting or escalating failings.	Identifies challenges to privacy and dignity and shapes their own care and that of others to ensure these are maintained. Takes time to find out more about their patient's needs to appreciate cultural and religious variations to privacy in order to provide individualised care.
7	The student demonstrates openness, trustworthiness and integrity.	Not perceived reliable to communicate or document accurate accounts of events. Does not keep promises. Not being open in admitting mistakes made or blames others. Failure to admit own limitations in experience, knowledge or skills.	Truthful and keeps promises. Delivers the care that has been committed to, keeps clear, timely records and communicates/ handover care or events accurately. Trusted to carry out patient care and honest about limitations; openly discusses mistakes and accepts responsibility for any short comings.	Exhibits professional conduct that reflects awareness of being a role model in representing the profession in all that they say or do, and in giving accurate accounts and documentation of events. Takes responsibility for their own actions and reports mistakes promptly and openly to minimise harm and is an advocate for such behaviour.
8	The student makes consistent effort to engage in and reflect on the requisite standards of evidence based care and learning to enhance care and their own professional development.	No consideration for evidence/ rationale/ research to support practice. Does not seek out information that will enhance care or own learning. Lacks interest or recognition of responsibility for learning and development. No efforts to reflect on personal and professional development.	Has a questioning and challenging approach to poor practice and keen to utilise research and local/national policy to establish evidence underpinning care delivery to enhance patient care and safety. Reflects on their practice to identify personal learning needs, and engages in learning/practice development and seeks out learning opportunities.	Able to search databases/journals and critically appraises the quality of evidence before applying it to their practice. Appropriately challenges poor practice and provides alternative solutions or points to evidence base. Up to date with local/national agendas or directives and acts as a resource to colleagues. Identifies practices that lack an evidence base and works with others to address this.

Cont. Figure 5.7: The final draft scoring rubric

5.7 Chapter Summary

This chapter has provided a detailed description of phase one of the research study which employed two rounds of Delphi to design a scoring rubric that provided level descriptors for the professional attitude, behaviour and responsibility statements in the PLPAD. Interpretations of three level descriptors for each of the eight statements were successfully developed in the first round. A strong stakeholders' consensus (ranging from 86%-100%) on these interpretations was achieved in the second round, eliminating the need for further rounds.

By using a Delphi method to design and develop a consensus-based scoring rubric in response to the need to improve mentors' effectiveness and confidence in assessment decisions, a novel and unique contribution to the field of pre-registration nursing education has been made. This study contributes to the body of knowledge by being the first to use a Delphi method to develop a scoring rubric for the professional attitude, behaviour and responsibility competency statements not only within nursing but across all practice-based assessment disciplines. As previously stated, the aim is to make practice-based assessment of pre-registration nursing students fit for purpose.

The following chapter outlines the second phase of the study. The research method used to evaluate the usefulness of the consensus-based scoring rubric in enhancing the quality of the formative and summative assessment of nursing students from mentors' and students' perspectives will be discussed.

Chapter 6: Evaluation of the scoring rubric

6.1 Introduction

The previous chapter presented the first phase of the study which involved two rounds of Delphi to design a consensus-based scoring rubric for the professional attitude, behaviour and responsibility competency statements in the PLPAD. This chapter outlines the method used for the second phase of the study, which evaluated the effectiveness of the scoring rubric in providing clear language and level descriptors, strengthening the rigour of mentors' assessment, and enhancing learning, self-assessment and feedback provision.

As discussed in section 4.2, the method is in keeping with the critical pragmatist world view that ideas are anticipations of possible solutions, and the *practical instrumentality* of the proposed alternative variations must be evaluated by the operation of acting upon them (Dewey, 1966). The quantitative and qualitative data were analysed and presented separately in this chapter as each element has its own analysis method. A combined analysis integrating both elements is provided in the discussion chapter (chapter 7).

6.2 Study design

This phase of the study adopted a convergent design; that is, the quantitative and qualitative methods were implemented concurrently during the same phase of the research process for the purpose of comparing or combining the results (Creswell and Plano Clark, 2017). In this type of design, both elements are analysed independently and then mixed during the interpretation stage (the design, including justification for its use, was discussed in section 4.3).

Evaluation research was the method employed to examine the effectiveness of the scoring rubric. Evaluation research, as a form of applied research, aims to produce information about the implementation, operation and ultimately effectiveness of a programme designed to bring about change (Moule *et al.*, 2017). Greater detail,

including justification for selecting this method as the most suitable approach for phase two in this study, was presented in section 4.4.3.

Two online questionnaires were used (a mentor and a student version) to collect quantitative and qualitative data concurrently. Based on the same rationale provided in the Delphi phase (section 5.4), the BOS was used as the website host to design and administer the questionnaires.

As described in section 5.5.2, the outcome from the pilot in phase one identified that attempting to interpret all the fourteen professional values statements was likely to affect negatively the response rate due to the length of time it took to complete the questionnaire. Consequently, in consultation with the supervisory team and the participants in the pilot, the decision was made to only use the professional attitude, behaviour and responsibility competency statements (n=8) and develop them into a scoring rubric. This meant modifying the original plan of testing the effectiveness of the scoring rubric experimentally by having two comparison groups, one using the scoring rubric (experimental group) and the other using the traditional process (control group). Instead both mentors and student were asked to complete the eight competency statements using the scoring rubric and to use the traditional process for the assessing the remaining six competencies that relate to 'safe and compassionate care'.

The modification from the original plan gave mentors and students the opportunity to experience using both the scoring rubric and the traditional process, which provided the advantage of reducing the amount of error arising from natural variance between individuals. Hence, according to Moule *et al.* (2017), fewer participants would be required, reducing the resources needed and recruitment demands.

Although all participants received the different exposures (assessment with the scoring rubric and without it), this may not fit a 'crossover design' completely, mainly because crossover designs are usually associated with providing exposure

sequentially rather than concurrently (Bowling, 2014; Gray *et al.*, 2017). Conversely, according to Shuttleworth (2009), the fact that both students and mentors received all the different exposures within the same assessment period, is potentially a crossover design since all the subjects receive all the exposures within the same study.

A commonly reported problem in crossover design is 'carryover effects'. This means that participation in one condition may affect performance in other conditions (Parahoo, 2014). For example, mentors may be tempted to use the scoring rubric to assess competency statements that are intended to be assessed through the traditional method. This was not considered to be an issue in this study for two reasons. First, the purpose of developing the scoring rubric was to produce level descriptors, specific and unique to each of the statements, making it non-transferable to other statements. Second, the professional values statements in the PLPAD are divided into two sections: the professional attitude, behaviour and responsibility section and the safe and compassionate care section (Appendix 1). The scoring rubric was designed for the level descriptors in the first section only, further limiting the carryover effects due to the differences in focus of the competencies.

6.3 Sample/Sampling

The study took a purposive sampling approach. Potential participants were full time third-year baccalaureate nursing students from the adult field within one London HEI located across two campuses. Sampling also included the students' named mentors across various clinical settings in London. Since the PLPAD competency statements differ for the first and second year students and for those who are studying other nursing fields, they were excluded due to the scoring rubric being specifically designed to interpret the competency statements in the PLPAD for the third-year adult nursing students only. However, it could be argued that due to the fact that all HEIs in London are using the same PLPAD and that the students were allocated to clinical placements spreading across London covering several acute and community NHS Trusts, it is plausible to argue that, although the recruitment of

students was from one HEI, it is potentially representative of all pre-registration nursing students using the PLPAD in London.

Mentors were approached following their identification as the allocated named mentor for those students willing to participate; this meant that they could be from clinical placements within a wide variety of placement providers contracted with the HEI. The criteria for selecting mentors were that they should hold a mentoring qualification approved by the NMC and meet the requirements to be registered as 'live mentors'. Most importantly, for them to be included in the study, both the student and their named mentor had to be willing to participate. The full inclusion and exclusion criteria for both students and mentors are provided in Table 6.1.

Table 6.1: Inclusion and exclusion recruitment criteria in phase two.		
	Inclusion	Exclusion
Students	<ul style="list-style-type: none"> • Third-year pre-registration • Adult field • Attended placement during the study • Consent to participate 	<ul style="list-style-type: none"> • First or second year students • Other fields (mental health, child, learning disability) • Named mentor not participating
Mentors	<ul style="list-style-type: none"> • Hold an NMC-approved mentoring qualification • Live mentor on the Trust register • Consent to participate. 	<ul style="list-style-type: none"> • Student not consenting • Absent during the placement

Onwuegbuzie and Collins (2007) discussed sample size in mixed methods and suggested that they tend to be dichotomised with large samples being associated with quantitative research and small samples with qualitative research. They argue that sample size should be informed primarily by the research objectives, questions and design. The main considerations for mixed methods sampling, according to Teddlie and Yu (2007), is the differentiation between probability and purposive sampling strategies. They explained that probability sampling (where the samples are gathered in a process that gives all the individuals in the population equal chances of being selected) leads to greater breadth of information from a larger number of units selected to be representative of the population. In contrast, purposive sampling leads to greater depth of information from a smaller number of carefully selected cases.

Teddlie and Yu (2007) recommended 30 cases or less for purposive sampling and a larger sample size of at least 50 to establish representativeness. In this study, although recruitment was purposive, efforts were made to enhance representativeness of the results, therefore, it was decided to recruit 30-50 pairs of participants. This number is also considered a reasonable sample size to enable interpretation of results in mixed methods (Onwuegbuzie and Collins, 2007; Creswell and Plano Clark, 2017).

The students' cohort within the local HEI under study comprised 190 third-year students across two campuses and all were considered eligible to participate. Before an invitation email was sent to students, the aim and objectives of the study were explained to them while attending sessions prior to their clinical placements. Once students who expressed interest to participate were identified (n=90), their clinical placement allocation was accessed from the placements allocation administrator at the HEI to identify the clinical placement area for each student.

To invite mentors, the researcher contacted the clinical areas to identify the allocated named mentors and then arranged to meet them in the clinical settings with their students within the first week of the placement to brief them together on the study including how the scoring rubric should be used. To be eligible for recruitment, both students and the named mentors had to agree to participate. To insure the researcher did not introduce researcher bias, the briefing session was standardised, using the participant information sheet as a guide. Some participants asked for more clarification than others; this was dealt with through rewarding the same information.

Out of the ninety students who expressed interest, fifty-one students (and their mentors) were recruited. Forty-eight pairs were recruited from eight hospitals within four NHS Trusts spreading geographically across north, south and east of London and three pairs from the community (see Table 6.2 overleaf).

Table 6.2: The sources of sampling for students and mentors in phase two.		
Placement	Potential participants	Recruited
Trust 1	18	7
Trust 2	9	3
Trust 3	16	7
Trust 4	40	31
Community	7	3
Total	90 (pairs)	51 (pairs)

6.4 Instruments

When deciding to select an appropriate tool for data collection in phase two, focus groups are considered ideal for discussing perceptions and experiences but they become less practical and less informative with larger number of participants (Krueger and Casey, 2014). The decision was therefore made to use questionnaires to gather individual mentors' and students' views after using the scoring rubric.

Questionnaires that concurrently collect quantitative and qualitative data are considered appropriate to measure opinions and attitudes (Oppenheim, 1992; Gray *et al.*, 2017), as well as allowing participants to raise issues they feel are important (Bowling, 2014). The development of the questionnaires incorporated a multistage process involving scoping the literature for validated questionnaires, question generation, and piloting.

The literature was reviewed to search for any validated questionnaires but there were no existing questionnaires found that matched the study objectives. Such a tool would have saved time and resources and enabled the findings to be compared with other studies. Instead, two online questionnaire versions (mentor and student versions) were specifically designed for this study using the BOS website as a platform.

6.4.1 Questionnaires design

There are different styles that can be used when developing a questionnaire to produce different types of data. Thus, it was important to be clear which scale and response format to use as this influences the analysis options (Rattray and Jones, 2007). The type of data in this study was ordinal as the variables ranked on a scale

of increasing magnitude and Likert scales are commonly used to collect this type of data (Wood *et al.*, 2011). Oppenheim (1992) emphasised that what makes Likert scales the most popular and appropriate scale to use, lies with their ‘unidimensionality’ making sure that all the items in a questionnaire measure the same thing.

A central step in planning a questionnaire is listing the topics of interest in relation to the objectives of the study (Bowling, 2014). Accordingly, the items generated in the questionnaires were guided by the research objectives in phase two (see Table 6.3). This also included grouping the items together by subject, and linking sentences were introduced when moving to a new group of questions (Oppenheim, 1992).

Table 6.3: Items used in the questionnaires and their relation to the research objectives and the findings from the scoping exercise and the IR.		
Research objectives	Findings from the scoping exercise and the IR	Items generated in the questionnaires
To evaluate the effectiveness of the consensus-based scoring rubric in providing clear language and level descriptors.	<u>From the scoping exercise (chapter 2, section 2.3)</u> Understanding the practice document/competence and the use of performance criteria. Identifying levels of competence.	<u>Mentor version</u> The rubric provided clear language to describe the levels. The rubric allowed me to distinguish between the different levels of competence. The rubric helped me recognise the required level the student needs to achieve. The rubric helped me explain to the student what is expected of him/her. The rubric reduced my confusion around the required level to pass.
	<u>From the IR (chapter 3)</u> Understanding the assessment process and criteria of the grading scheme. Distinguish levels and description to reflect student performance accurately. Understanding the language used and consistency in interpreting and defining what should be assessed in the criteria.	<u>Student version</u> The rubric helped me recognise the mentor’s expectations of the required level I need to achieve. The rubric provided clear language of the described level. The rubric allowed me to distinguish between the different levels of competence. The rubric increases understanding around the required level to pass. The rubric provided adequate direction to help me understand the required level of attitudes and behaviours.

Table 6.3 cont.: Items used in the questionnaires and their relation to the research objectives and the findings from the scoping exercise and the IR.		
Research objectives	Findings from the scoping exercise and the IR	Items generated in the questionnaires
To evaluate the effectiveness of the consensus-based scoring rubric in strengthening the rigour of mentors' assessment.	<p><u>From the scoping exercise</u> What needs to be demonstrated to be worthy of a pass. Mentor interpretation of student ability. Uncertainty (process/knowledge of programme/student needs). Time.</p> <p><u>From the IR (chapter 3)</u> Guidance to the amount of supporting evidence needed to award levels. Confidence in the decision made about the student performance. The mentor-student quality of relationship and communications. Directions/objectivity to help in assessing attitudes and behaviours.</p>	<p><u>Mentor version</u> The rubric provided adequate direction to me to understand and assess the student attitude and behaviour. The rubric provided a clear guide to measure student performance accurately. The rubric gave me confidence that the decision and grading of the student is fair. The use of rubric will help me to be consistent in identifying and grading students' level of performance in the future. The rubric will ensure all mentors involved in student assessment are consistent in accurately identifying the required level of competence.</p> <p><u>Student version</u> The rubric allowed me to see the fairness of the decisions made to my level of competence. The rubric will ensure all mentors involved in student assessment will be consistent in accurately identifying the required level of competence.</p>
To evaluate the effectiveness of the consensus-based scoring rubric in enhancing learning, self-assessment and feedback provision.	<p><u>From the scoping exercise</u> Guiding/influencing reflection on practice (real life, authentic).</p> <p><u>From the IR (chapter 3)</u> Ability to give accurate feedback to support grade and clarify learning.</p>	<p><u>Mentor version</u> The rubric provided a structured approach to identify and plan the student's learning needs. The rubric facilitated accurate feedback that supported the level awarded. The rubric guided me to reflect on practice and real-life situations with the student. The rubric was a useful tool to encourage reflective discussion between me and the student about his/her professional attitudes and behaviour.</p> <p><u>Student version</u> The rubric provided a structured approach to identify and plan my learning needs The rubric allowed me to self-assess my own level. The rubric guided me to reflect on practice and real-life situations with my mentor. The rubric was a useful tool to encourage reflective discussion between me and the mentor about my professional attitudes and behaviour.</p>

Collecting demographic data is considered a useful way to split respondents into groups to examine how the groups vary with their responses (Bowling, 2014). Chapter one identified that the efficacy of mentors' assessment of nursing students is multifactorial, and analysing factors related to variations within groups may yield valuable information. However, the aim in this phase of the study was specific to evaluating mentors' and students' opinions on the effectiveness of using the scoring rubric. Therefore, collecting demographic data was not considered essential in order to meet the study aim and objectives, hence, there was no justification for collecting such data. In addition, the advantages of not including a category about demographic data may have reduced the attrition rate by not asking personal questions and making the questionnaires shorter (Oppenheim, 1992).

The order and structure of items

General recommendations about questionnaire development and presentation were followed including providing clear instructions, appropriate length, ease of read, question numbering, avoiding splitting questions between two pages and providing enough space for free-text comments (Oppenheim, 1992; Rattray and Jones, 2007; Bowling, 2014). The neutral category was included in the questionnaires based on the same rationale as when the second questionnaire for phase one was developed (see section 5.4).

Considerations were also given to the presentation order of items. The 'funnel approach', which involves starting off with broad questions and progressively narrowing them down to the specific issues (Bowling, 2014), was used. The questionnaires started with questions about the experience of using the scoring rubric as a tool in general, such as their ability to recognise the performance levels and their understanding of the language, and then asked increasingly specific questions about the degree of accuracy and fairness of their decisions. The most important questions, such as those related to the rigour of the assessment when using the scoring rubric, were asked first to ensure that important data was not entirely lost if the questionnaires was not fully completed (Bowling, 2014).

The filtering approach, where respondents are directed to the next applicable question (Oppenheim, 1992; Bowling, 2014) was not thought to be needed in this study as all items were applicable. Bowling (2014) advocated avoiding filtering questions if possible as they create real potential for errors.

Wording of the questions

The language and wording of questions can affect responses; hence attention was paid to using short, simple and familiar words and phrases that virtually all respondents would be able to understand and conceptualise in the same way (Bowling, 2014). Leading or complex questions (including double negative and double-barrelled questions) were also avoided (Oppenheim, 1992; Rattray and Jones, 2007; Bowling, 2014). The questionnaires were also piloted on a representative sample (see next section) to further examine the wording, readability and content of the questionnaires.

‘Acquiescent response bias’, where respondents tend to agree with a statement, or respond in the same way to items, and ‘response set’ known as the tendency to endorse the responses positioned at one end (Bowling, 2014), were considered. Rattray and Jones (2007) argued that acquiescent response bias and response set can be minimised by using a mixture of both positively and negatively worded items and alternating the type and direction of response codes.

However, the decision was made not to use positively and negatively worded items or alternate directions. This was based on studies reporting that respondents will still use a ‘response set’ and reply to all the items as if they were positively worded, concluding that negatively worded questions bring about inconsistency and confusion (Colosi, 2005), and that fewer mistakes are made if items are posed in the same direction (Sonderen *et al.*, 2013). Such unfavourable consequences were experienced in the responses to the questionnaire in the Delphi phase (discussed in section 5.6). Therefore, the decision was made to keep all items positively worded in the questionnaires.

'Satisficing' is also an order bias that occurs when attitude statements are presented in a list. Satisficers will read until they find an adequate answer that satisfactorily and sufficiently reflects the answer, rather than reading the full list (Brace, 2008). This type of bias is likely to increase with long questionnaires as respondents stop making efforts to answer to the best of their ability, therefore, attempts were made to only include items relevant to the research questions to keep the questionnaires as short as possible and minimise satisficing.

Social desirability bias is another threat to validity where respondents choose 'prestigious' answers. Oppenheim (1992) stated that this is common in attitude and opinion questionnaires making accurate interpretations difficult. Oppenheim (1992) advised reiterating to respondents that accuracy is the prime requirement and that negative responses are just as valid as positive responses. Bowling (2014) also suggested using natural words that reflect everyday speech to enhance reporting of socially undesirable opinions (e.g. using 'Do you think...?' rather than 'In your opinion...?'). However, since the questionnaires in this study was asking respondents to share their experience of using a novel approach to assessment, it was considered that the effect of social desirability was not significant since all the possible answers were neutral due to the focus being on how useful the scoring rubric was rather than the participants' own aptitude. Therefore, no answer would be considered more prestigious than the other.

Oppenheim (1992) advised that hypothetical questions, whereby respondents are asked to make predictions about the future are best avoided as they are found to be poor predictors of people's future reaction or behaviour. However, hypothetical questions that related to consistency by asking both mentors and students if they thought the scoring rubric would help mentors to be consistent in future assessments were included in the questionnaires (see questions 15 and 12, Appendix 18 and 19 respectively). Since consistency cannot be accurately tested until mentors assess several students using the scoring rubric, it was thought a hypothetical question may provide some insight. Nonetheless, this was an example where the neutral option would prove useful.

As mentioned earlier, two questionnaires were designed, a mentor version (Appendix 18) and a student version (Appendix 19). The first part of each questionnaire comprised Likert scale questions with fixed choice responses ranging from strongly agree to strongly disagree allowing descriptive statistics to be used to evaluate their experience of using the scoring rubric. The second part had open-ended questions to give participants the opportunity to expand on their answers and provide more in-depth responses.

6.4.2 Methodological rigour

In this study, the intended purpose of the questionnaires was to evaluate the usefulness of the scoring rubric; hence steps to establish validity were focussed on achieving face and content validity. Face validity is defined by Gray *et al.* (2017, p. 678) as “a subjective assessment by an expert to verify that an instrument appears to measure the content it is purported to measure.” This was achieved by piloting the questionnaires with an expert group comprising nurse lecturers and practice educators to ensure that the questionnaires were appropriate for the study purpose and content area.

Content validity is defined as “the extent to which the measurement method includes all the major elements relevant to the construct being measured” (Gray *et al.*, 2017, p. 674). This was achieved when the questionnaires’ items (and the research objectives) were designed by amalgamating the data collected from the scoping exercise and the findings from the IR (presented in Table 6.3). Additionally, content validity was augmented by submitting the questionnaires to six experts (two academics, two practice educators and two PhD supervisors) who reviewed items for potential ambiguity of the wording and to identify any unnoticed flaws in any of the items.

Lastly, in the validation of an instrument, Timmerman *et al.* (2011) stated that tools should be evaluated in terms of the intended purpose of the tool. Thus, content validity was also derived from piloting the questionnaires on a sample representing the target groups to make certain that participants interpreted questions and

response categories as intended. This was to elicit comments on the clarity, format and ordering of questions, and to evaluate whether the intended end users believed the scoring rubric asked the right questions (Coughlin *et al.*, 2009).

Reliability of the questionnaires, referred to by Gray *et al.* (2017, p. 690) as “the consistency of the measures obtained”, examines the number of measurement errors in the instrument being used and is usually expressed in correlation coefficient based on the aspect of reliability being examined. The main three aspects of reliability according to Gray *et al.* (2017) are stability (the consistency of repeated measures over time by measuring test-retest reliability), equivalence (the consistency between scores by measuring inter-rater reliability), and internal consistency (the homogeneity of all the items by measuring Cronbach’s alpha coefficient).

The aim for developing the questionnaires was specific to this study and intended to gather mentors’ and student’ perspectives after they had used the scoring rubric. Hence, the approach taken to enhance reliability and reduce variations focussed on maintaining methodological rigour when designing the questionnaires as discussed earlier. In addition, all the students and their allocated mentors were invited to participate to ensure selection bias and researcher influence were minimised. Reliability was also enhanced through rater training, when both mentors and students were briefed on how the scoring rubric should be used throughout the placement.

6.5 Procedure

This section describes how the evaluation research method was conducted, including piloting, data collection and data analysis. This phase of the study took place from 14th November 2016 to 17th February 2017.

6.5.1 Ethical Considerations

This phase of the study was given a separate approval by the School of Health and Social Care Research Ethics Panel at the HEI where the research was conducted (Appendix 20). Permission from the Head of Department at the HEI was obtained prior to contacting the students (Appendix 21). To access mentors, the first step was to use the NHS Research Ethics Committee online tool to check if the study required approval; this indicated that it did not need NHS approval (Appendix 22). However, approval from the individual NHS Trusts' education leads was obtained, which, for some of them, was subject to their local ethics clearance (Appendix 23). An indemnity letter from the HEI (Appendix 24) was supplied to Trusts on request.

Separate participant information sheets for students (Appendix 25) and mentors (Appendix 26) were given to participants when they were briefed on how to use the scoring rubric. They also had the opportunity to ask questions. Consent was also obtained in writing (Appendix 27) during the briefing session, assuring them about the anonymity and confidentiality of data collection and storage, and reiterating that participation was voluntary and that they were free to withdraw from the study at any time without giving any reason. However, they were also informed that data would no longer be retrievable once it had been anonymised and merged with other data.

The researcher did not have direct teaching responsibilities with any of the participants that could cause a conflict of interest, and contact was limited to the agreed communication channels for the purpose of the study. No personal information was shared including not giving out private phone numbers and ignoring any social media requests that might have been sent by participants.

Consideration was given to the fact that students may feel anxious about using the scoring rubric as an adjunct to their practice assessment document, fearing that it may disadvantage them. Such effect was thought to be minimal mainly because the competency statements in the PLPAD were unchanged in the scoring rubric. The only thing the scoring rubric added was the provision of level interpretations for the

statements. In an effort to mitigate potential anxiety, the researcher addressed the issue in both the information sheet and the briefing sessions, and participants were reminded that they could withdraw from the study at any point without the need for explanation. Contact details for further guidance and support were provided. However, it is acknowledged that this issue may have been a factor in recruiting students.

Consideration was also given to anonymity and confidentiality. In research, anonymity refers to participants not being identifiable to others (Farrimond, 2013). True anonymity could not be promised as both mentors and students needed to be identified and briefed on using the scoring rubric. Furthermore, identity was disclosed between each student and the named mentor. However, data were collected anonymously through the BOS, which is a secure website where only the invited participants had access. All responses were strictly confidential, and participants' identities were not disclosed. Direct quotes to free-text answers used as part of the study report were not traceable. However, it was acknowledged that there may be a chance that participants talked among themselves about the study.

Confidentiality in research is concerned with not sharing data about the participants beyond what has been consented for (Farrimond, 2013). To allow analysis, responses from the online questionnaires were downloaded and stored on a password protected computer using participants identifying numbers only. Participants were made aware that they had the right to access submitted information according to the UK data protection laws, but that due to anonymity, identifying individual responses would not be possible.

6.5.2 Piloting the online questionnaires

Prior to the commencement of data collection, the questionnaires were piloted on a small sample comprising eight mentors and eight students matching the characteristics of the target sample but not included in the main sample. They were given a copy of the scoring rubric and the link to the online questionnaires. They were asked to comment on accessibility, length, clarity, time it took to complete,

the layout, give suggestions for alternative wording, raise objections to answering any questions and if they had any other comments. Their comments did not yield any issue that needed amendment, and on average it took them 5-10 minutes to complete the questionnaire.

6.5.3 Conducting the study

When the students started their placement, the researcher arranged to meet the students and mentors in the clinical settings to explain what was expected of them by taking part in the study. This included briefing them on how the scoring rubric should be used throughout the placement and to obtain their written consent and an email address to send them the questionnaires. A copy of the scoring rubric (presented in Figure 5.7) and the participant information sheets (Appendix 25 and 26) were given to participants. The scoring rubric was administered over an eight-week placement from 14th November 2016 to 20th January 2017.

The day after the clinical placement was completed, both mentors and students were sent an email with a link to the online questionnaires. The website started with a welcoming screen that included the researcher contact details. This was followed by a series of screens for the closed-ended questions, and a final screen comprising four open-ended questions. The questionnaires ended by thanking the participants for their contribution (copies of the questionnaires can be found in Appendix 18 and 19). The website was made available for three weeks from 23rd January 2017 to 17th February 2017 with reminder emails sent after one week (Appendix 28) and a final reminder one week later (Appendix 29).

6.5.4 Data analysis (closed-ended questions)

Data from the Likert scale were ordinal and since there was no comparison between groups, descriptive statistics were the most appropriate approach. Responses were analysed using the BOS website as data were already presented with a tally of response totals and percentage of responses for each statement.

There are substantial variations in how the results of descriptive statistics are reported in research (Thabane and Akhtar-Danesh, 2008), with evidence that the quality of it is sub-optimal (Latronico *et al.*, 2002; Mills *et al.*, 2004). Thabane and Akhtar-Danesh (2008) reviewed the available guidelines for reporting studies arguing that consideration for the characteristics of the sample or data is essential and provides a description of the process to determine what descriptive statistics to use and how to report the results. They recommended that data from categorical or ordinal variables should be reported as numbers with the percentages next to it in brackets, claiming that most papers use this approach to summarise the results.

Greene (2008) also supported this conclusion adding that percentages provide an important tool to illustrate nurses' opinions of the set of pre-determined questions. However, there is a debate that if the sample size is too small it will not yield a meaningful percentage. Lang and Secic (2006) specified that in a sample size less than 20, the actual number should be reported rather than percentages, as in small samples percentages can be misleading because the size of the percentage is so much greater than the number it represents. Thus, it was decided that using numbers and percentages in brackets would provide clear and adequate reporting of the quantitative data from the Likert scales. It was also decided that calculating the mean, median, mode, range and standard deviation would not add significant value.

6.5.5 Data analysis (open-ended questions)

The qualitative findings from the free-text comments were transferred into tables and the analysis process included reading the comments to identify key messages. Due to the relatively small amount of data collected, as most of the comments were a one-line sentence, it was decided that content analysis (that looks for repeated ideas or patterns of thoughts to categorise the key messages) would be more appropriate than thematic analysis which is more suited to extract latent meaning from a phenomenon of interest, commonly working with large amounts of data to find overriding abstract ideas (Gray *et al.*, 2017).

Similar to the analysis in the first round of Delphi, a deductive content analysis approach following the same steps of the content analysis employed in the Delphi (see section 5.5.4) was used. The units of analysis were selected from the frequently repeated comments, which comprised the units of meaning from the statements that had the same meaning. Special considerations were given to map the units of analysis to the research objectives and the closed-ended questions to facilitate integration of the quantitative and qualitative data (Denzin, 1970). After reading through the comments several times to become familiar with the data, a categorisation matrix was then developed, and the key messages were coded according to their category (Appendix 30).

To enhance face validity, the content analysis was independently verified by the supervisors. In qualitative research, verification refers to the mechanisms used during the process of research to incrementally contribute to ensuring reliability and validity and, thus, the rigour of a study (Creswell, 1997). Morse *et al.* (2002) argued that qualitative researchers should reclaim responsibility for reliability and validity by implementing verification strategies from external reviewers during the conduct of their inquiry to ensure rigour.

6.6 Results

In this section, the results of the participants' perceptions on the effectiveness of the scoring rubric to answer the research aim and objectives are presented. The first section presents the results from the closed-ended questions and the second section presents the findings from the open-ended questions. Analysis and integration of both these sets of results is provided in the discussion chapter (chapter 7).

Response rate

Of the 51 pairs of mentors and students recruited, 43 (84%) mentors and 46 (90%) students responded to the online questionnaire. On reviewing the data, there were a few instances of contradictions between the closed-ended and open-ended responses (3 mentors and 1 student). For example, in the closed-ended questions,

some of the respondents strongly disagreed with the statement that they found the scoring rubric was overall useful but provided very positive comments in the open-ended questions when they were asked if they found it useful.

According to May (2010), mixed methods research is prone to contradictions in data because of the different categories and levels of analysis and dealing with such data depends on whether the purpose behind data integration in mixed methods research is triangulation or complementary. Denzin (1970) argued that the purpose for a triangulation approach is that the findings from one method can be validated by using another method. Denzin stressed that findings are expected to converge in order to be validated therefore, contradicting findings are problematic and cannot co-exist in this approach. In contrast, in the complementary approach, methods do not investigate the same object but rather connect segments of social phenomenon that complement one another to construct social reality (Mason, 2006). Under this approach, conflicting findings represent different viewpoints on the same phenomenon and the fact that social reality is complex (May, 2010).

This argument guided the decision to exclude contradicting data in this study as the overall purpose of collecting quantitative and qualitative data in the questionnaires was to investigate the same objective in order to produce a fuller picture. However, it is worth noting that based on the strong results obtained (presented below), exclusion of these responses would have minimal effect on the overall results. In addition, one mentor was excluded due to submitting an empty questionnaire and another questionnaire was excluded as the student stated in the comments section that the scoring rubric was not used. Therefore, the final presentation of the results is based on responses from 39 mentors and 44 students.

There were some respondents who strongly disagreed that the scoring rubric was useful, but there were no comments provided by participants in the free-text box to explain the reasons. It would have been useful to know why they strongly disagreed, but due to the anonymity of responses it was not possible to follow them up for clarification.

In the open-ended section, there were four questions and not all those who returned their questionnaires completed this section.

- The first question asked what they like best about the scoring rubric and 37 (84%) students and 30 (77%) mentors responded.
- The second question asked what they least liked about the scoring rubric, and 35 (80%) students and 21(54%) mentors responded.
- The third question asked if they found the scoring rubric overall useful with 38 (86%) students and 30 (77%) mentors responding.
- The final question provided them with the opportunity to add any comments; 21 (48%) students and 14 (36%) mentors responded.

6.6.1 Results from the closed-ended questions

This section used closed-ended questions and sought participants' perceptions on the effectiveness of the scoring rubric. The results are presented in a sequence related to the study's objectives:

Objective 1: To evaluate the effectiveness of the consensus-based scoring rubric in providing clear language and level descriptors.

Both mentors and students were asked if the scoring rubric helped them to recognise the required level of competence needed to be achieved. Overall 34 of the 39 mentors (87%) and 43 of the 44 students (98%) strongly agreed or agreed that the scoring rubric helped them distinguish the required level (Figure 6.1).

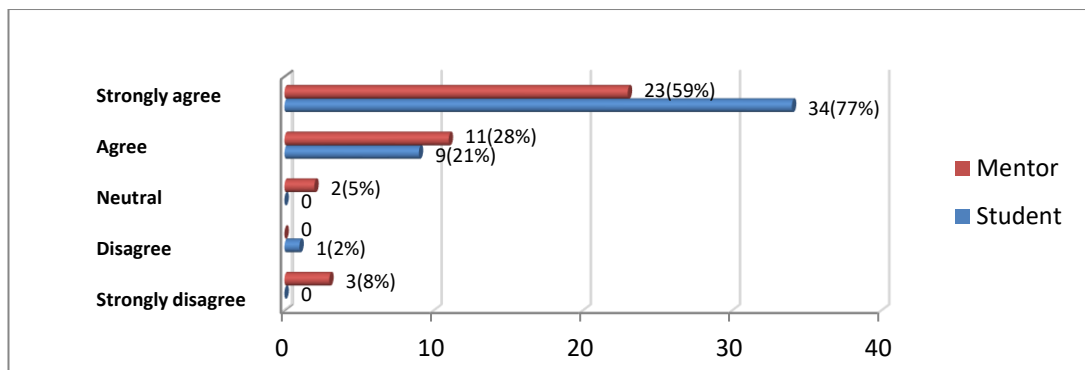


Figure 6.1: Recognising the required level to achieve

Mentors' and students' views on the clarity of language that describe the levels were also sought. The majority of mentors and students found that the language used to describe the level descriptors for each competency statement was clear (Figure 6.2).

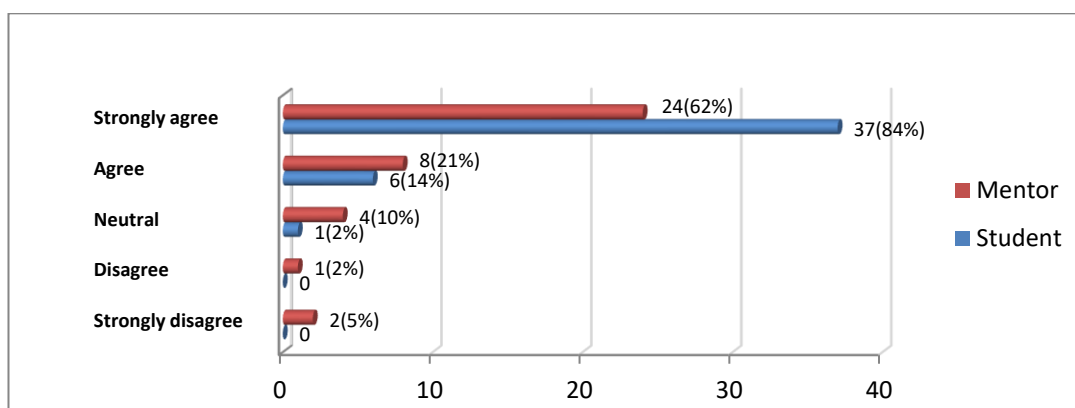


Figure 6.2: Clarity of language describing the required levels

A question that was only directed at the mentors asked if the scoring rubric helped them to explain to the students what was expected of them. The majority of mentors (85%) strongly agreed or agreed that the scoring rubric helped to explain what was expected of students (n=33) (Figure 6.3).

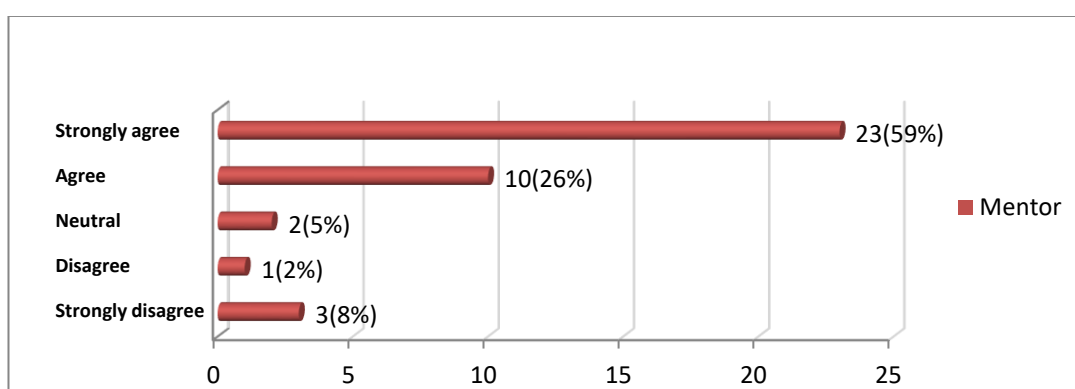


Figure 6.3: Help to explain what is expected of students

Both mentors and students were asked if the level descriptors in the scoring rubric ('Does not meet expectations', 'Meets expectations' and 'Exceeds expectations') distinguished the three levels of competence. The majority of mentors and students

thought that the descriptors provided clear distinction between the three levels of competence (Figure 6.4).

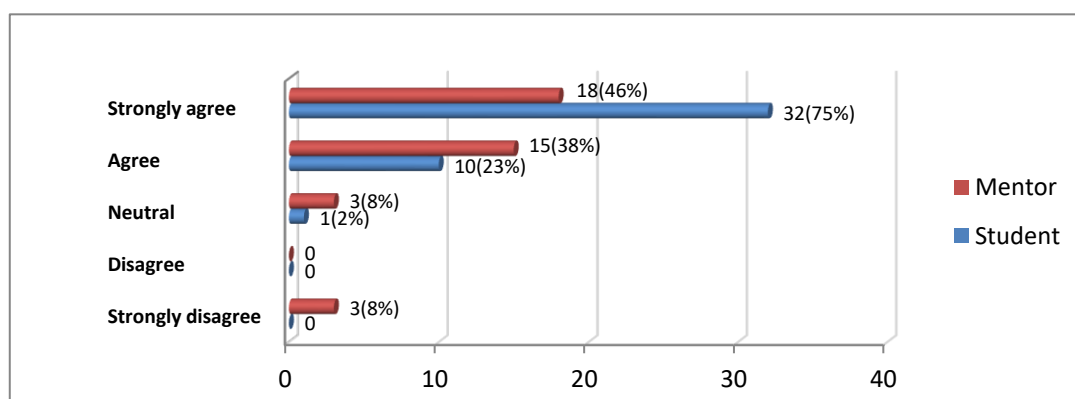


Figure 6.4: Distinguishing between different levels of competence

When asked if the scoring rubric helped them to understand the required level to pass, 34 out of 39 mentors (87%) strongly agreed or agreed that the scoring rubric reduced their confusion around the required level to pass. Out of 44 students, 42 (96%) also strongly agreed or agreed that it increased their understanding of the required level to pass (Figure 6.5).

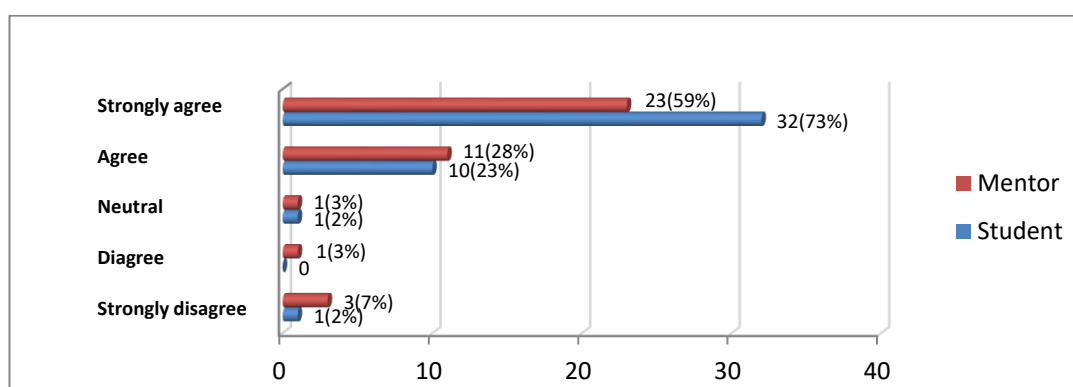


Figure 6.5: Understanding the required level to pass

Participants were also asked if the scoring rubric provided adequate direction to understand and assess the professional attitudes and behaviours. Of the 43 students who responded to this question, 42 (98%) thought the scoring rubric provided adequate directions to understand the levels of competence for the professional attitudes and behaviours. Likewise, 84% of mentors (n=33) reported that they strongly agree or agree that the scoring rubric provided adequate

directions to understand and assess professional attitudes and behaviours (Figure 6.6).

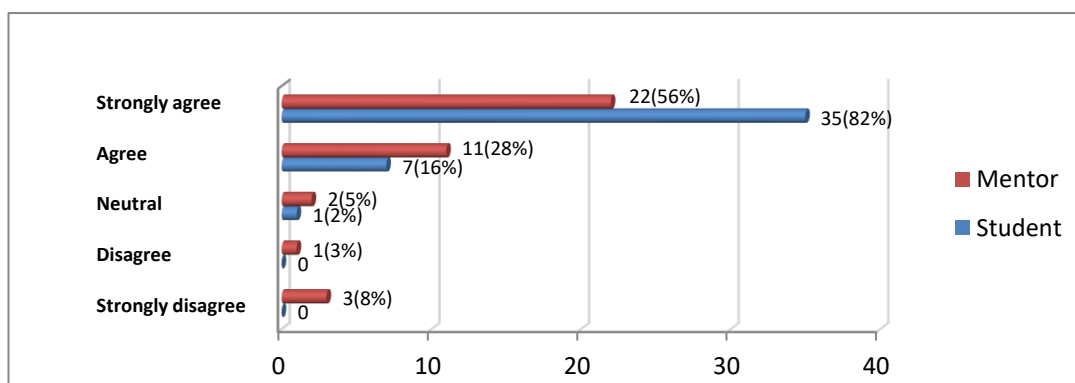


Figure 6.6: Directions to understand and assess attitudes and behaviours

Objective 2: To evaluate the effectiveness of the consensus-based scoring rubric in strengthening the rigour of mentors' assessment.

Both mentors' and students' perspectives on the usefulness of the scoring rubric in reaching fairer assessment decisions were sought. Of the 39 mentors, 33 (85%) strongly agreed or agreed that the scoring rubric gave them confidence that the decision and grading of the student was fair, and 41 of the 44 students (95%) strongly agreed or agreed that the scoring rubric allowed them to see the fairness of the decisions made to grade their level of performance (Figure 6.7).

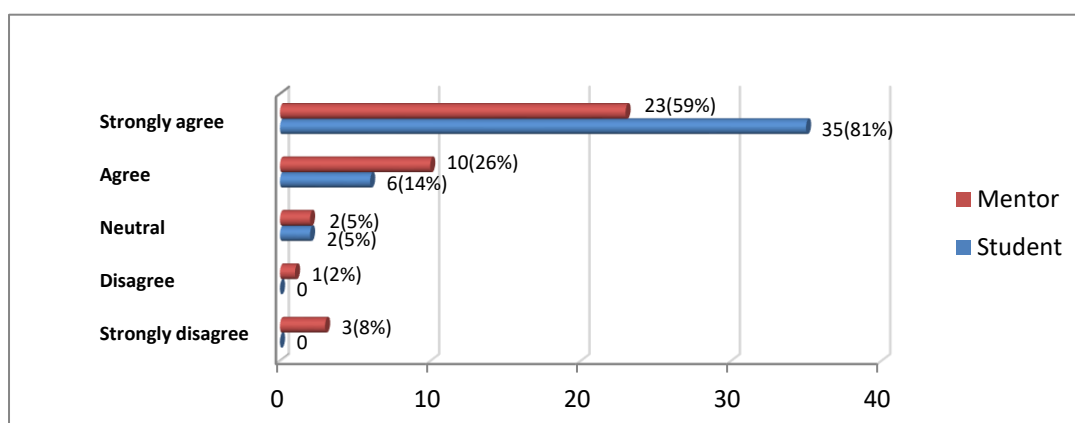


Figure 6.7: Fairness of grading decisions

A question aimed at the mentors only, asked them if the scoring rubric improved the accuracy of measuring student's performance. Of the 39 mentors, 31 (82%)

indicated that the scoring rubric provided them with a clear guide to measure student performance accurately (Figure 6.8).

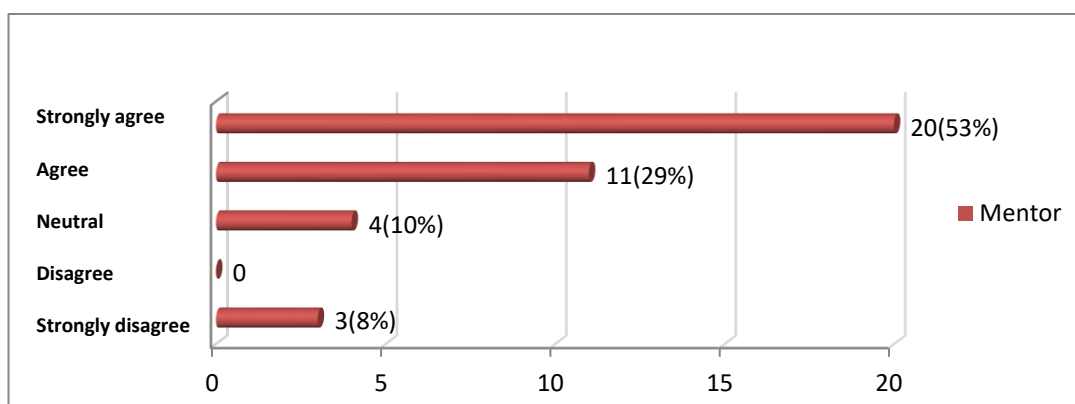


Figure 6.8: Measuring student's performance accurately

To explore the ability of the scoring rubric to produce reliable and consistent results, mentors' perceptions of their own practice were sought first, and 34 of the 39 mentors (88%) strongly agreed or agreed that the scoring rubric will help them to be consistent in identifying and grading students' level of performance in the future (Figure 6.9).

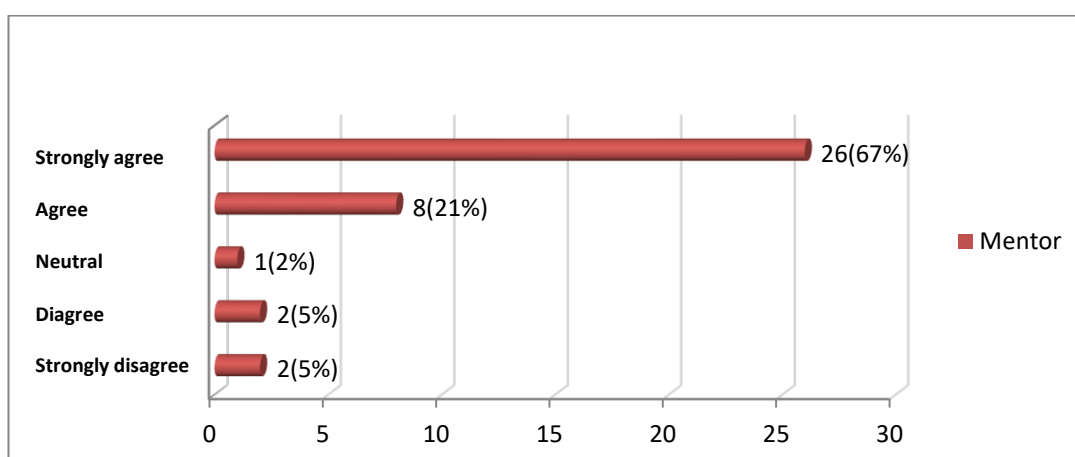


Figure 6.9: Mentors outlook on their own consistency

In addition, both mentors' and students' views about consistency in general were sought; 33 out of 39 mentors (85%) and 43 out of 44 students (98%) strongly agreed or agreed that the use of the scoring rubric will ensure mentors involved in

assessing students will be consistent in accurately identifying the required level of performance in the future (Figure 6.10).

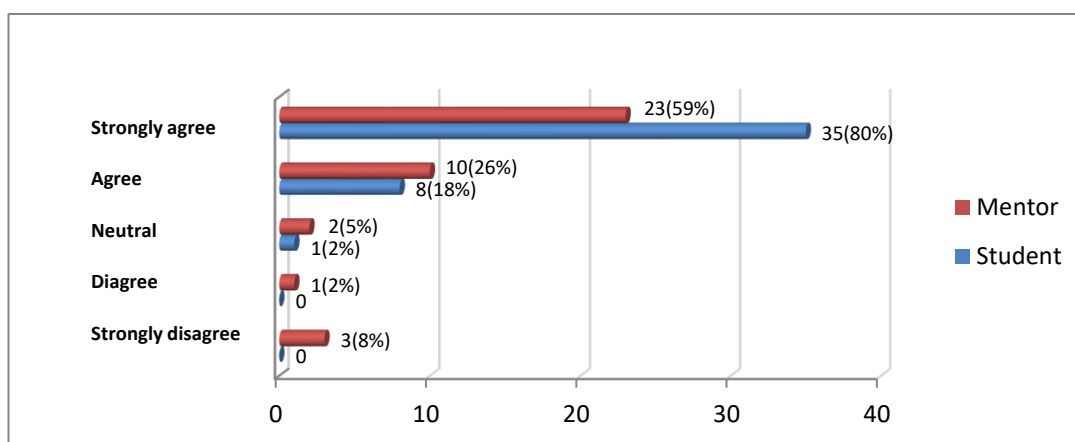


Figure 6.10: General outlook on the rubric consistency

Objective 3: To evaluate the effectiveness of the scoring rubric in enhancing learning, self-assessment and feedback provision.

To evaluate the usefulness of the scoring rubric in facilitating learning and feedback, the majority of mentors and students reported that the scoring rubric provided a structured approach for the identification and planning of learning needs (Figure 6.11).

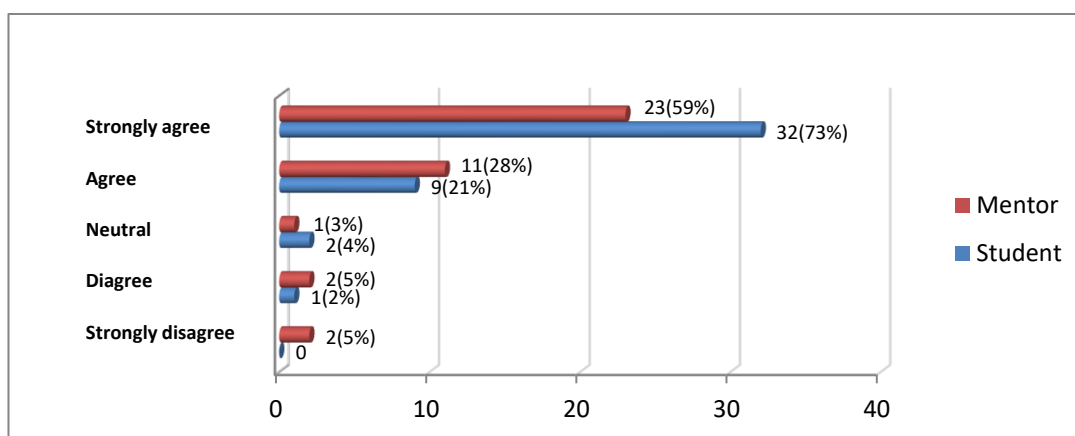


Figure 6.11: Structured approach for identifying learning needs

In terms of the usefulness of the scoring rubric for mentors to give feedback that supports assessment decisions, 34 mentors (87%) indicated that they strongly agree

or agree that the scoring rubric helped them to give accurate feedback that supported the level awarded (Figure 6.12).

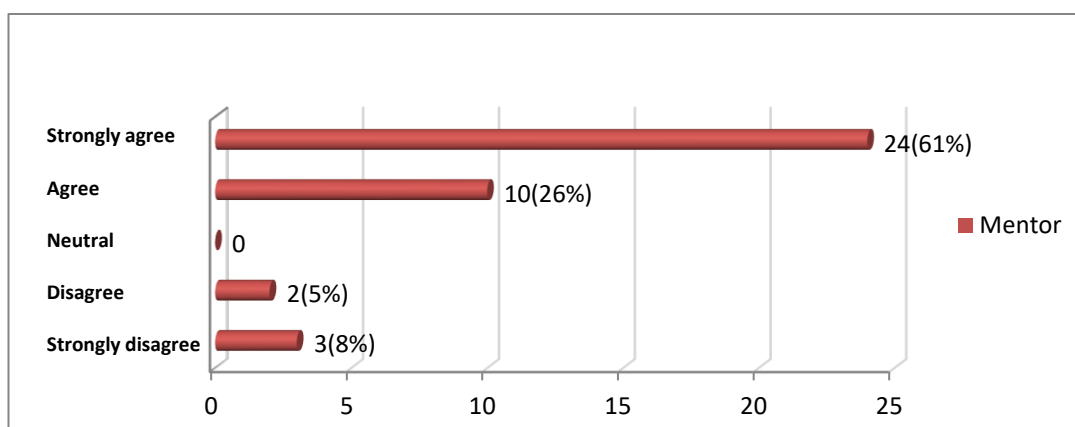


Figure 6.12: Giving accurate feedback to justify the level awarded

The question about the usefulness of the scoring rubric in allowing self-assessment was aimed at the students and 42 out of 44 students (95%) reported that they strongly agreed or agreed that the scoring rubric helped them to identify their own level of performance (Figure 6.13).

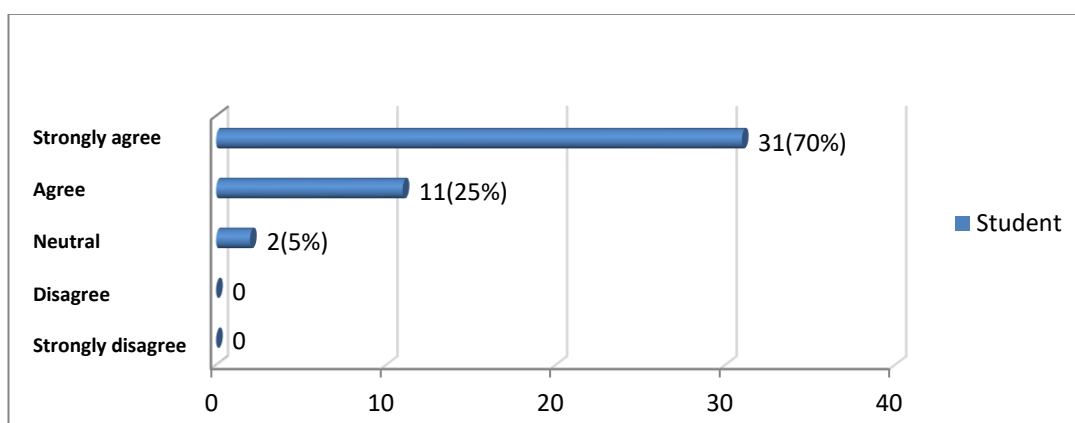


Figure 6.13: Students self-assessment

When asked if the scoring rubric encouraged reflective learning, 32 mentors (82%) and 39 students (89%) strongly agreed or agreed that the scoring rubric guided reflection on practice and real-life situations (Figure 6.14).

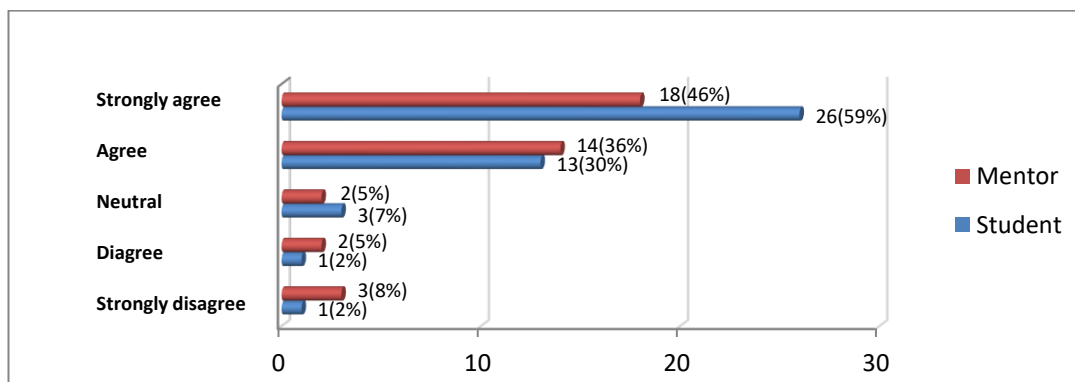


Figure 6.14: Reflection on practice and real-life situations

The majority of mentors (n=30) and students (n=40) indicated that the scoring rubric encouraged reflective discussions between them about professional attitudes and behaviours (Figure 6.15).

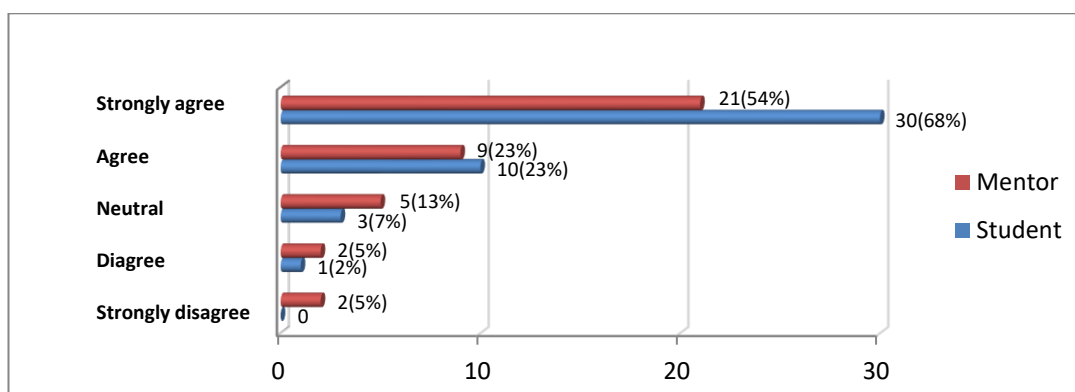


Figure 6.15: Reflective discussion about professional attitudes and behaviours

The final question asked the participants about the practicality of the scoring rubric, 34 of the mentors (88%) and 42 students (96%) strongly agreed or agreed that they found the scoring rubric practical to use (Figure 6.16).

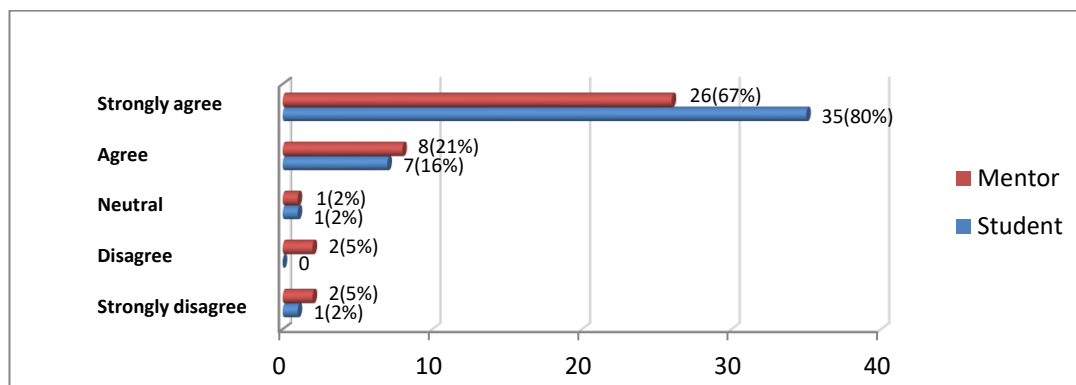


Figure 6.16: Practicality of using the rubric

6.6.2 Findings from the open-ended questions

This section presents the findings from the open-ended questions to provide a fuller picture on the effectiveness of the scoring rubric in the assessment process from both mentors' and students' perspectives. The qualitative findings add *authenticity* to complement the quantitative results. The analysis process to reach categories of the key messages was discussed in section 6.5.5, the quotes provided were selected to illustrate the key messages (Appendix 30 presents all comments). The shared views of mentors and students were amalgamated into the following eight categories:

i. Clarity of the language

Both mentors and students reported many positive comments about the clarity and structure of the scoring rubric. They felt that the scoring rubric offered simplicity, guidance and structure that helped them understand the assessment document. Several mentors described how their improved understanding enhanced their knowledge of the assessment process.

“Easy to understand and made scoring of student more straightforward and was easy to explain to student.” Mentor 7

"It was easy to use, and much clarity in reducing any conflicts or misunderstanding." Mentor 23

The mentors also commented positively on the clear language especially that it gave them the appropriate language to use in their communication.

"The words to put up while writing about the student objective, what to look for when an objective is asked, very useful." Mentor 2

"Very informative in helping to choose the right word in completing students pack." Mentor 4

Encouraging comments from students and mentors were also related to the clarity and simplicity of the language used in the scoring rubric. From the students' perspective, they seemed to appreciate the well-defined and unambiguous language used.

"Both me and my mentor found it useful as so many times in the past the wording of the PAD itself was daunting." Student 9

"Sometimes the concepts in the PAD can be abstract and vague and it helped give more clarification." Student 24

Interestingly, students also appeared to recognise mentors' increased knowledge and understanding.

"It was easier for my mentor to assess me by having clear and fair guidelines." Student 5

"It gives clarity of what is expected of you. Hence helps both the student and mentors." Student 40

ii. Distinction between levels of competence

The usefulness of the scoring rubric as a tool to determine different levels of achievements was also reported by mentors and students. Mentors commented on how the scoring rubric provided an insight into what is expected of the students and how to grade them.

"Found it very useful as a guidance to distinguish between 'fail/pass/and exceptional performance.'" Mentor 19

"...I also liked that it allowed me to say whether they were achieving expectation or merely meeting it - much better than just saying achieved or not achieved." Mentor 32

The students shared the same views of the mentors and they were very positive in their comments. They found the scoring rubric very useful during their placement by making it easier to understand different levels of outcomes through explaining the expected level.

"It helped give a more extensive explanation of what was expected of me." Student 24

"Sets out what exactly I need to do to achieve my assessments." Student 22

Students also appreciated that the scoring rubric recognised those who go beyond the required level of achievement.

"I think that the idea of fail, pass or exceeds is a great way to determine the difference between students who pass or excel during placement." Student 16

Students also reported that the scoring rubric helped identification of the required attitudes and behaviours in the professional values.

"It helped me a lot to understand the required behaviours and attitudes." Student 17

"Explained the professional values thoroughly making it easier to distinguish a pass and a fail." Student 30

"It was very precise about what was expected of me unlike before where the values seemed vague." Student 37

A noteworthy observation in the students' comments was their recognition that having a mutual understanding was ultimately crucial in the assessment process.

"The rubric was very useful in guiding the mentor and as a student provided guidance as to the standards we should be reaching." Student 2

"It gives clarity of what is expected of you. Hence helps both the student and mentors." Student 40

iii. Rigour of the decision-making

Many respondents referred to the usefulness of the scoring rubric in maintaining rigorous and fair decision-making process. Mentors described how the scoring rubric was useful for *"evidencing assessment"* (Mentor 37), and *"decision-making is easier"* (Mentor 39). A key finding was that the scoring rubric reduced the subjectivity of their assessment of students.

"It gives clear and concise view and plan so one does not assess a student on assumption. It also assists the facilitator in making their decisions." Mentor 28

"It gave me a clearer insight into what was expected of the student and how to grade them" Mentor 22

The usefulness of the scoring rubric in enhancing consistency of the assessment was also an important finding by several mentors.

"I think some facilitators can be very rigid in their opinion about students; this will give a uniform plan to assess students." Mentor 28

"Helping the students understanding and your own understanding of the level needed for passing the placements...more standardised." Mentor 31

The students' comments were even more powerful in stating the usefulness of the scoring rubric to the rigour of the assessment. Their satisfaction with the fairness of assessment was a strong finding.

"I liked most about the rubric style of assessment is the fairness...and gives me confidence that I will be judged entirely fairly." Student 22

"It allowed fair and fast assessment." Student 23

"Made the midpoints and finals more structured and more thought put into the marking." Student 4

The students also shared the mentors' views about the objectivity and consistency benefits of using the scoring rubric. This is a very important finding.

"It eliminates any worries that I will be judged on anything but my performance." Student 22

"Made the decision process smooth and more accurate." Student 23

"This process allows mentors to assess students solely on their nursing skills." Student 38

"The best thing is that it gets rid of any bias or unachievable expectations by a mentor." Student 22

iv. Facilitation of learning, feedback and self-assessment

The usefulness of the scoring rubric as a tool to support learning and the provision of structured and constructive feedback was recognised by both mentors and students. Mentors found the scoring rubric helpful in providing instructions to advance learning and was instrumental in making action plans not only to manage poor performance, but also to push students who have higher potential to excel.

"The rubric allowed me to discuss in detail what was required of the student to achieve higher levels in each area...It also allowed you to build an action plan to achieve excellence and not just when students are failing to meet standards expected." Mentor 32

The students' responses about the value of the scoring rubric as a learning tool were very powerful. Many students described how useful it was in enhancing learning and that it explained what is required of them with clear directions on how to achieve them.

"Helps to have something to work towards." Student 30

"It gave me a sense of direction." Student 21

"It provided me with a standard of practice that I could recognise and build upon if necessary." Student 44

"Allowed myself to comprehend the required elements to achieve each component of my professional values." Student 39

"Really defines for students HOW they can achieve their professional values. Much better than the current statements which only tell you WHAT they are expecting." Student 15

Some of the comments indicated that the scoring rubric went beyond just helping them to achieve what is required showing that students shared mentors' views that the scoring rubric encouraged them to flourish.

"The rubric provided me with a target to exceed placement instead of just passing." Student 16

"...Also allows for professional development and recognises behaviours which are better than simply pass." Student 30

"It provided guidelines for assessment and a target to achieve the exceptional status." Student 31

An additional advantage reported by the students was the usefulness of the scoring rubric in promoting self-assessment and reflection on practice.

"Helped me to understand what my strengths and weaknesses are...also helped me to assess my own progress." Student 36

"It makes it much easier to see what level you are at and how to improve or maintain your standards." Student 21

"Allowing me to use the rubric to reflect upon situations I have encountered and dealt with in my practice." Student 39

v. Time

An important finding was the usefulness of the scoring rubric in reducing the time spent documenting the assessment process, making the interviews shorter.

"Saved time in filling out paper work." Mentor 19

"It helped with the flow of the students assessment and the interview was shorter which helps when you are busy." Mentor 2

Mentor also suggested time was saved because the scoring rubric provided clarity

"Save time and easy to understand." Mentor 15

Even a student recognised that the time factor helped mentors.

"Saves time for mentors." Student 15

Students' views on the other hand were more mixed. Some students shared mentors' views that the scoring rubric saved time,

"It allowed fair and fast assessment." Student 2

Other students reported that the scoring rubric prolonged the assessment process.

"It's more time consuming especially when you're on a busy unit."

Student 21

"It did make the process more time consuming but I don't know how that could be avoided." Student 24

Interestingly, the time factor was reported positively by students, and they discussed how it enhanced their learning.

"Slightly more time consuming but allowed for a deeper discussion into what exactly the professional value is." Student 30

"Even though it's time consuming, it makes it much easier to see what level you are at and how to improve or maintain your standards." Student 21

vi. Student-mentor interaction and dialogue

The relationship and communication between mentors and students runs throughout the responses, with the scoring rubric appearing to give both mentors and students a shared purpose. This mutual understanding was an important factor in building professional relationships that foster growth and promote learning and development. Several mentors shared this view.

"Prompted good discussion between myself and the student."
Mentor 11

"I was able to communicate and reflect with the student on her entire placement." Mentor 23

The students were also very positive in their comments about the influence of the scoring rubric on the mentor-student relationship and interaction. It is worth noting that these are third year students and speak from a position of knowledge and experience.

"This time round it wasn't a mentor giving you their version of what they think that means but working together with you to achieve what both of us know should be achieved. Awesome really." Student 34

"Provides a framework to guide to discussing and achieving professional values." Student 31

"It gave the mentor a guide...and helped me to understand what she thought my strengths and weaknesses are." Student 36

A further advantage reported by the students was related to the usefulness of the scoring rubric in avoiding conflict situations.

"Prevents conflicts and disagreements between mentors and students." Student 5

"Very useful for some students who may experience personality clashes." Student 38

"I can imagine a situation in which the mentor might doubt someone is achieving but the student would have evidence to demonstrate that they have achieved the expected requirements." Student 36

vii. Participants reflections on using the scoring rubric

There were some reflections provided by both mentors and students about their experience using the scoring rubric. One mentor commented that the scoring rubric *"still needs perfection"* (Mentor 6) but did not provide any further elaboration. A comment from one student expressed concerns of being more examined and observed.

"I felt nervous as if every part of me was going to be scrutinised more because of how specific it was." Student 37

Another student felt that the 'exceeds expectations' descriptor was set at a higher standard that should be expected of students.

"The 'exceeds expectations' category were higher than what even the average nurse performs at." Student 36

Several mentors and students reflected positively about the scoring rubric and recommended its use. Students comments suggested that the scoring rubric gave them confidence and courage and it was empowering.

"I am glad to have used the rubric." Student 10

"I think it would be fantastic for everyone to use....and I will recommend it." Student 36

"Very useful I hope it comes into use and is used for all students." Mentor 27

"Would be useful if implemented into practice for students in the future." Student 40

"I hope it will be used for all students to give them the confidence, and structure they need." Student 22

"I think it would be a wonderful tool that would empower and encourage all students." Student 22

Both mentors and students expressing their desires of wanting a scoring rubric for all the competency statements.

"Not a finished complete tool yet...would love to see the finished result." Mentor 31

"Would be more helpful on the other values as they are more wishy-washy." Mentor 37

"It was only for the first eight questions. Would have preferred the rubric for all the questions." Student 8

"Would like rubric for all the competencies in future." Student 9

6.7 Chapter Summary

This chapter provided the method and results of phase two of the study where evaluation research was used to establish the effectiveness of the scoring rubric from mentors' and students' perspectives when they have used it in practice.

Both quantitative and qualitative data were collected and analysed, with important results and findings indicating that the scoring rubric did enhance the quality of the formative and summative assessment of nursing students. There is clear evidence that both mentors and students found the scoring rubric helped them understand the assessment document. It provided clear and unambiguous language thereby reducing the risk of different interpretations. Distinction between achievement levels was also facilitated by the scoring rubric and resulted in identifying what are the expectations and how to grade them.

One of the most valuable findings in the study was the usefulness of the scoring rubric in improving reliability and validity of the assessment, ensuring the assessment process is rigorous, fair and consistent. The usefulness of the scoring rubric was also demonstrated in facilitating learning, constructive feedback and self-assessment.

Some positive findings emerged from the qualitative data that were not directly explicit in the study objectives. These included, enhancing mentor-student interaction/dialogue and saving mentors' time.

The next chapter (chapter 7) provides a detailed discussion incorporating how the results from the quantitative and qualitative elements were integrated to combine inferences and provide a fuller, contextualised perspective by reference to published literature, and the theoretical frameworks (assessment *for* learning and authentic assessment) that underpinned this study.

Chapter 7: Discussion

The previous chapter presented the method and results of the second phase of the study to evaluate the effectiveness of the consensus-based scoring rubric. The aim of this chapter is a) to provide a synthesis of the quantitative results and qualitative findings of the evaluation phase of the study, b) to examine the most valuable of these in relation to the impact of the scoring rubric on making practice-based assessment fit for purpose and c) to compare and contrast the findings with the evidence examined in the IR (chapter 3) as well as relevant studies that have been published since completion of the IR. In keeping with the critical pragmatism principle, which calls for ideas to be designed and tested for their *practice instrumentality* (Dewey, 1966), the results of this research study will be critically examined to establish the extent to which the designed scoring rubric enhanced practice-based assessment.

The chapter will also review whether the theoretical frameworks of ‘assessment *for* learning’ and ‘authentic assessment’ used to underpin this study enlighten understanding of the findings. Alternative theoretical frameworks that may help illuminate the processes underlying using the scoring rubric in practice-based assessment will also be examined.

7.1 Synthesis and contextualisation of findings

Assessing the practice element of the pre-registration nursing curriculum against competencies set by professional bodies is essential to evaluate that learners have developed an adequate level of competence, ensuring that they are safe and fit for practice in order to protect the public (Yorke, 2005; Trede and Smith, 2012). However, the literature reviewed in chapter 2 established that many studies raised concerns regarding lack of reliability and validity of assessment in practice placements, concluding that it is not fit for purpose.

The literature reviewed in this study (chapters 1 and 2) identified that there are numerous barriers to effective assessment of the practice element, including

unfamiliarity with the assessment document, with particular difficulties in understanding the language used or in identifying the appropriate level of competence. This potentially compromises the reliability and validity of practice-based assessment, which ultimately could result in unsafe registrants. As explained in chapter 3, a well-designed scoring rubric with transparent and unambiguous language to interpret different levels of competence was proposed as a possible solution to these challenges. This could be achieved through helping mentors define what is expected of students and for students to identify what they are expected to achieve (Jonsson and Svingby, 2007; Gray and Donaldson, 2009; Shipman *et al.*, 2012; Frentsos, 2013). Consequently, the overall aim of this study was to develop a scoring rubric and evaluate if interpreting level descriptors for the professional attitude, behaviour and responsibility statements in the PLPAD make practice-based assessment of pre-registration nursing students fit for purpose.

Several themes emerged from the findings, including some that were not directly explicit in the study aim and objectives (see section 3.6), which will also be analysed in this section. These relate to mentor-student interaction/dialogue and saving mentors time.

7.1.1 Clarity of language and distinction between levels of competence

The IR revealed that within the context of assessing competence in nursing, a gap in interpreting and differentiating between performance levels existed and contributed to lack of rigorous assessment which is essential for patient safety (see section 3.3.4). Therefore, the effectiveness of the scoring rubric in making practice-based assessment fit for purpose will be discussed here in the context of clarity of language and the distinction between levels of competence.

i. Clarity of language

Evaluating the effectiveness of the scoring rubric in providing mentors and students with clear language to help understand the terminology of the assessment document was an objective for the second phase of the study. Both the IR (chapter 3) and the scoping exercise conducted locally (section 2.3) confirmed that there are

difficulties in the language used to describe competencies in practice assessment documents. Competency statements were described as vague, open to interpretation and difficult to translate (Neary, 2001; Dolan, 2003; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Cassidy *et al.*, 2012). This resulted in the frequently reported problem of subjectivity in assessing competencies in practice and stressing the need for clear and unambiguous language (Dolan, 2003; Butler *et al.*, 2011; Cassidy *et al.*, 2012). The literature also revealed that current assessment tools provide generic descriptors that lack specificity so remain open to interpretation (Neary, 2001).

Findings from this current study illustrate that the consensus-based scoring rubric helped both mentors and students to recognise and describe the level needed to be achieved and that the language to describe the levels was clear. Additionally, the quantitative and qualitative responses confirm that the scoring rubric, with its matrix of clearly described levels of expectations articulated in simple language, was effective in improving understanding and implementation of the assessment process. Mentors found the clarity of terminology useful in allowing them to explain to students what was expected of them.

The scoring rubric succeeded in addressing the ambiguity of terminology problems, reported frequently in assessment tools (Brown, 2000; Duffy and Watson, 2001; Norman *et al.*, 2002; Cassidy *et al.*, 2012; Helminen *et al.*, 2016), and echoed in a recent Australian study evaluating the way nurses understand their competency standards. In that study, Terry (2017) revealed that the majority of nurses did not understand the language in the standards nor how they are connected to the assessment of competence, thereby leaving the reliability and validity of assessments largely dependent on the subjectivity of the assessor's interpretation.

Similarly, a recent study in ROI by Burke *et al.* (2017) found that mentors experienced difficulties in understanding the content of the assessment document, reporting problems in negotiating the perceived complex language. In another recent study, exploring mentors' experiences of assessing student nurses, Cassidy *et*

al. (2017) found that the continued problem of failing to fail is driven by mentors' struggle to interpret practice learning outcomes appropriately and convey the meaning of those issues to others. These recent studies support the findings reported in the IR in chapter 3 (McCarthy and Murphy, 2008; Butler *et al.*, 2011; Cassidy *et al.*, 2012). These three additional studies affirm that ambiguity of the terminology used in assessment documents continues to be a global problem in nursing and one for which this study provides a solution.

In addition, although previous literature reported that ambiguous language is problematic, this study uncovered for the first time what actually concerns mentors when the language is not clear. The qualitative responses illustrated that poorly understood language is a source of conflict and misunderstanding. Lack of clarity also contributed to difficulties in articulating, communicating or documenting what was required from students. Mentors in this study reported that the scoring rubric not only assisted them in accurately explaining to students what is required of them to achieve, it also helped them write appropriate language when completing the assessment document. The usefulness in documentation was frequently reported in relation to choosing the right language when there was a need to articulate an action plan and to communicate mentor decisions. What this study also revealed is that the students not only mirrored the mentors' comments, they also recognised that their mentors had an improved understanding which made the assessment process easier, clearer and fairer.

The overall shared message that emerged from the responses is the lack of clarity in the current format of the PLPAD, which brings into question the reliability of the assessment. As discussed in section 1.1, this lack of clarity is not unique to the PLPAD but characterises practice assessment documents worldwide across a variety of professions, not just nursing. The threat that unreliable assessment poses to public protection has been stated frequently throughout this thesis, but looking at this problem from a different angle, the lack of mutual understanding may also mean that students are not seeing fairness in practice placements and more consideration is needed to 'protect the student'. It is likely that feeling of 'not

getting a fair deal' is doing students a disservice and may be a contributing factor in their level of dissatisfaction and attrition rates. This concern is worth investigation in future research.

ii. Distinction between levels of competence

Evaluating the effectiveness of the scoring rubric in allowing mentors and students to distinguish between performance levels of practice was also an objective in the second phase of this study. Both the quantitative and qualitative responses from students and mentors demonstrated that the scoring rubric enabled them to recognise performances at different levels and to identify the appropriate level students need to achieve. This is an important finding since both the scoping exercise and the IR established that mentors struggle to discriminate between different levels of practice or identify the benchmark of what constitutes a pass or a fail (Girod, 2000; Neary, 2001; McCarthy and Murphy, 2008; Butler *et al.*, 2011; Fahy *et al.*, 2011; Heaslip and Scammell, 2012). This struggle includes identifying the required professional attitudes and behaviours, reported in the literature to be the most difficult domain to assess (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016).

An updated review of the literature demonstrates that this problem continues to be reported globally. Terry (2017) studied nurses' utilisation of the Australian National Competency Standards for Registered Nurses; her findings echoed the IR conclusions that there were variable levels of expectations between academics, clinical facilitators and mentors in their understanding of the performance expected. In a similar study, Burden *et al.* (2018, p. 1087) found that mentors' frame of reference in judgement decisions is based on impressions directed by generic, pre-existing expectations to decide that a student is "safe enough to pass", posing the difficulty of not being compatible with programme and professional standards. Additionally, the responses from the qualitative comments in this study illustrated that students were ill-informed regarding what was expected of them in placements. Their comments signify, yet again, the importance of mutual understanding of what is expected to be achieved and the fairness of the process to

students. Therefore, this study shows that even if mentors are able to differentiate between levels of performance, if such performance levels are not shared and explained to students, the assessment process would be perceived as ineffective and unfair.

The qualitative comments from both mentors and students also revealed the advantages of having more than just a dichotomy, such as the achieved/not achieved classification currently adopted in the PLPAD, or other commonly used variations such as 'competent/not competent' or 'pass/fail'. Providing three distinguishable levels of performance that allowed mentors to identify at which level students were working was valuable in this study.

This explanation is reinforced by Burden *et al.* (2018) who reported that mentors found it difficult to determine their decision when proficiencies were assessed using only achieved/not achieved threshold, suggesting that the use of grading may assist mentors' decisions and improve their ability to establish the standard of practice achieved. Grealish and Show (2018) supported this argument when they reviewed The Australian Nursing Standards Assessment Tool (ANSAT) introduced nationally in 2016. Similar to the PLPAD's dichotomised classification, the ANSAT classifies students' performance as satisfactory (competent) or not (not competent). Grealish and Show (2018) debated that when assessors are confronted with the significant decision with the only choices being to pass or fail a student, they tend to award a pass due to the consequences of failing them. In addition to the 'failure to fail' phenomenon, this can wrongfully suggest that no further improvement is required. This led Grealish and Show (2018) to conclude that adding a category of 'not yet competent' is urgently needed, echoing Heaslip and Scammell's (2012) conclusion that assessment tools that use more discriminating grading systems were helpful and welcomed by mentors. This current study confirms that students also welcome discriminating grading systems.

Additional benefits were gained in this study through moving away from dichotomised classification towards more positive language, which is one of the

principles of scoring rubrics (Allen and Tenner, 2006). Using positive terms for the level descriptors to reflect that the student is 'working towards competence' rather than 'not competent', appeared to improve provision of constructive and criterion-based feedback for the struggling student. Equally, for those students who achieved competence, the scoring rubric provided developmental scope to excel and improve, illustrated by one of comments made by a mentor in this study that the scoring rubric is '*a great way to determine the difference between students who pass or excel*', which is something existing assessments fail to do.

Summative assessment was also facilitated by the scoring rubric. Instead of the judging of the student being left open to individual mentors' intuition and experience, and only limited to achieved/not achieved, the same scoring rubric is designed to be used formatively to give repeated exposure and practice opportunities to gauge the performance as well as increase mentors' familiarity with the outcomes and competencies being assessed at the summative point.

7.1.2 Strengthening the rigour of assessment

One of the most valuable findings in this study is that the scoring rubric enhanced the rigour of the assessment process. The comments from mentors revealed that the scoring rubric assisted them in making informed decisions about what was expected of the student as well as allowing them to justify their decisions. This finding is an important contribution to knowledge since previous literature only provides evidence that mentors struggle to justify their decisions or prove their concerns are valid, which are frequently reported as factors contributing to failing to fail underperforming students (Duffy, 2003; Gainsbury, 2010; Lawson, 2010; Brown *et al.*, 2012).

The current study also illustrates that the use of a scoring rubric further improved the rigour of practice-based assessment through enhancing consistency. The scoping exercise and the IR reiterated that lack of consistency in practice-based assessment is a major issue including inconsistency in how statements are interpreted (Dolan, 2003; Butler *et al.*, 2011; Fahy *et al.*, 2011). The IR also revealed

that mentors have differing views about what is considered as the 'acceptable' standard of competence a student needs to meet in order to pass, which compromises interrater reliability (Gray *et al.*, 2017).

Mentors in this study felt that the scoring rubric supported consistency in grading students and could see how it will help mentors involved in assessing students to be consistent in identifying the required level of performance. The qualitative comments reinforced this view as mentors in this study recognised how the scoring rubric made the assessment '*more standardised*' and provided uniformity of a shared mentor-student understanding. A view also shared by the students.

Achieving a shared mentor-student uniformity of understanding is an important finding, considering that Dolan (2003, p. 136) reported inconsistencies in the way students, mentors and university lecturers interpret competency statements, both within and between groups, adding that lecturers felt "students need to accept that mentors may have their own interpretations of competence". Thus, students experienced variations concerning what assessment criteria should be and felt challenged to figure out what was expected of them (Neary, 2001).

Recent literature also identifies that inconsistencies in student nurses' assessment processes vary not only between countries but also between institutions, with little evidence for strategies to overcome this problem (Helminen *et al.*, 2016). The detrimental effect of such variations on determining fitness for practice and public protection has been frequently reported in the literature (Yorke, 2005; Trede and Smith, 2012).

The open-ended responses also illustrated that the scoring rubric enhanced objectivity. Mentors revealed that the clear and precise language and descriptors in the scoring rubric allowed their decision, as one mentor reported, '*not to be made on assumptions*'. This finding reiterated that, currently, assessment of students in practice is based on impressions, supporting previous literature that mentors rely on the subjectivity of their intuition, first impression and gut instinct when making

decisions (Black, 2011; Burden *et al.*, 2018). Additionally, students felt the scoring rubric eliminated mentors' biases and that it gave them the confidence that they will be judged fairly and solely on their performance. This is a valuable finding as it indicates that the scoring rubric facilitated the development of trust, considered an important factor in building mentor-student relationships and understanding of each other's roles and expectations (Haitana and Bland, 2011). This resulted in increased student satisfaction with the fairness of the assessment, thereby meeting Hardy *et al.*'s (2015) expectation that students should have a positive clinical experience in general.

This study also revealed that the use of a scoring rubric with explicit descriptions of performance levels written in an easily understood language enhances the accuracy of practice-based assessment. Both the scoping exercise and the IR (see chapters 2 and 3) identified that available tools and taxonomies failed to adequately assess students' competence (Butler *et al.*, 2011; Fahy *et al.*, 2011; Heaslip and Scammell, 2012). A plausible explanation of the reason why the scoring rubric succeeded where other tools have previously failed is that the scoring rubric provided explicit and objective criteria. This contrasts with the current grading tools that provide generic descriptors lacking in specificity, which remain open to differing interpretations (Norman *et al.*, 2002; Watson *et al.*, 2002; Duffy, 2003; Moore, 2005; Cassidy *et al.*, 2012).

This explanation is also reinforced by Calman *et al.*'s (2002) and McCarthy and Murphy's (2008) arguments that, despite previous tools being based on taxonomies, mainly Benner's (1984) novice-to-expert, such tools have been criticised as being unworkable, with problems related to difficulty in understanding the wording or determining the level of competence. This results in mentors not finding them useful and therefore not using them during the assessment process (Neary, 2001). It is worth noting that more recent evidence continues to criticise previous tools used in assessing performance describing them as being "blunt instruments" (Burden *et al.*, 2018, p. 9). The evaluation of the scoring rubric in this current study shows it overcomes the problems of these previous tools.

The complexity of assessment in the inherently unpredictable real world (Govaerts and Van der Vleuten, 2013), compounded by the fact that professional attitudes and behaviours are notoriously challenging to define or measure (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016), needs acknowledging. This study is unique in showing that a scoring rubric enhanced accurate assessment of the professional attitude, behaviour and responsibility competency statements in the PLPAD, and gave mentors confidence that their judgement of a student's level of competence was fair. Additionally, this study illustrates that the students were able to see the fairness of the decisions made and they reported how the scoring rubric allowed comprehension of the professional values and provided a framework to guide discussing and achieving them. The increased accuracy of assessing attitudes and behaviours is a unique and important finding and may reduce the challenges mentors experience when making a judgement to determine whether the behaviour observed fulfils the required standards (Black *et al.*, 2014; Hunt, 2016).

7.1.3 Enhancement of learning and feedback provision

Another objective in the evaluation phase of this study was to establish the usefulness of the scoring rubric when used in the formative assessment to enhance accurate identification of performance levels and provision of constructive feedback to move the student forward as necessary. As discussed in the IR, scoring rubrics are theoretically underpinned by the concepts of assessment *for* learning and authentic assessment (Wiggins, 1990; Halonen *et al.*, 2003; Frentsos, 2013). This is based on the premise that scoring rubrics capture the qualitative aspects of meaningful learning by making expectations and criteria explicit and transparent, which facilitates feedback, self-assessment and communication (Jonsson and Svingby, 2007; Shipman *et al.*, 2012).

The findings of this study demonstrate that mentors found the scoring rubric helpful in giving accurate feedback. The qualitative comments revealed how instrumental the scoring rubric was when there was a need to provide structured and constructive feedback, especially when making action plans to manage poor

performance. The scoring rubric also helped mentors with their choice of language when documenting evidence. It is evident that clear and transparent criteria increased mentors' knowledge and understanding of different levels of competence and enhanced their assessment literacy. This allowed them to explain and articulate what previously had been based on intuition and was difficult to express in the documentation as reported in Duffy's (2003, p. 21) study where mentors failed to "put pen to paper". More recently, Burden *et al.* (2018) found that despite assessment processes being followed when mentors raise a concern about a student who is not meeting expectations, documenting evidence was sparse, and mentors generally provided verbal rather than written feedback. In her PhD thesis, Burden (2014) concluded that better documentation practices are needed and recommended further research into approaches that provide an adequate amount of information to illustrate transparent judgements and identification of students' developmental needs.

Therefore, this study not only reinforces the arguments presented in the IR, which emphasised the importance of formative assessments in providing individualised feedback on performance, it also reveals the advantages associated with the scoring rubric by using positive language. Thus, rather than highlighting incompetence, it offers encouraging and personalised information as to how the student can become competent or, as Schuwirth and van der Vleuten (2011) explained, how to steer student learning to reach full potential and strive for excellence. This also rightfully aligns assessment to the educational curriculum by placing the purpose of assessment upon learning and development (Biggs, 2003), which has been long called for in the literature (Ali, 2013; Govaerts and van der Vleuten, 2013; Haines *et al.*, 2013). Equally, this study shows that the use of grading (as opposed to pass/fail) augmented the assessment *for* learning principles by allowing identification of areas for learning and planning how to address them.

7.1.4 Student-mentor interaction and dialogue

Mentors and students recognised the importance of interaction and dialogue which, as clearly illustrated thus far, runs throughout the responses. The consensus-based

scoring rubric appeared to give a shared purpose that fostered growth and promoted learning and development.

Although communication in practice-based assessment has been frequently reported in the literature, it is viewed as a skill to be applied and assessed in practice (Shakespeare and Webb, 2008), with the focus mainly on maintaining effective communication skills with patient and colleagues (Pellat, 2006; Wilkes, 2006) or formal written communication both in documentation and feedback provision (Brown, 2000). This current study provides an interesting dimension about utilising communication as a tool to establish competence. The open-ended responses from both mentors and students showed that, when discussing professional attitudes and behaviours, the scoring rubric stimulated discussions and interactions to reach a joint identification of what constitutes 'competent'. This is illustrated by a statement from one student describing how the scoring rubric helped both the student and the mentor through '*working together to achieve what both of us know should be achieved*'.

This statement is powerful and carries two important messages. The first message is the importance of having *mutual* understanding of what is expected to be achieved. Thus, even if mentors were able to understand the language and differentiate between levels of performance, if such understanding is not shared with the student, the assessment process is ineffective. Earnshaw (1995) arguably would support such an explanation having identified that the relationship between mentor and mentee works best when both parties contribute equally to the relationship. The second, more substantial, message relates to utilising communication as a *source* of making judgement. The interaction and dialogue provided mentors with opportunities to engage with the students, allowing them to make sense of the student-communicated responses or dynamics and rationales.

It is worth noting that the imminent separation, as required in the newly published NMC standards (NMC, 2018a), of the roles of practice supervisor and practice assessor which cannot be performed by the same person, means it is anticipated

that students will spend most of their time with a range of practice supervisors and occasionally a practice assessor. This is likely to result in students spending less time with their assessor compared to the current practice, which can be detrimental to the development of mutual understanding between students and their assessors.

In contrast to the current practices discussed in chapter 2, where decisions about students' fitness for practice were based on subjective perceptions and intuition (Duffy, 2006; Black, 2011; Burden *et al.*, 2018), this study reveals that employing explicit and transparent criteria in the scoring rubric stimulates debate between mentors and students and inspires dynamic professional growth that fosters self-direction and critical thinking. More importantly, the scoring rubric provided mentors with more concrete insight and justification of students' competence. The study shows that assessors can emancipate their assessment practices towards a critical, wise and self-defining practice that nurtures lifelong learning in students and assessors alike (Trede and Smith, 2014), thereby embodying Dewey's (1966) critical pragmatist values of involvement, participation and openness.

7.1.5 Time

Lack of time to supervise and assess students effectively was the most frequently reported barrier in the scoping exercise (see section 2.3). This is consistent with a substantial amount of literature that identifies lack of time as a key and chronic issue compromising the mentors' role as 'gate keepers' (Atkins and Williams, 1995; Watson, 2000; Pulsford *et al.*, 2002; Duffy, 2003; Webb and Shakespeare, 2008; Robinson *et al.*, 2012; Veeramah, 2012; RCN, 2015; Hunt, 2016). As discussed in chapter 2, considering the overstretched and understaffed clinical areas (NHS Employers, 2014), and despite the DH (1999), the NMC (2008a, 2018e) and the RCN (2015) calls for protected time, enforcing such a change that includes freeing mentors from their clinical duties was considered very unlikely to be endorsed by the healthcare providers due to the shortage in nursing manpower as well as the associated financial implications. Hence making research inquiries and proposing ideas to improve time was deemed unrealistic to pursue, and accordingly was not

an objective in this study despite its presumed substantial impact on the practice-based assessment process.

Although participants in this study were not asked to comment about time, responses to the open-ended questions from both mentors and students stated that the scoring rubric saved time. The most plausible explanation to why it saved time is likely to be related to the clarity of language and level descriptors. Arguably, ambiguous statements in the practice assessment documents require mentors to spend time trying to make sense of what they mean. This explanation is supported by previous findings that both mentors and student spend significant amount of time trying to deconstruct the competency statements to fit with their practice rather than assessing the student against them (Neary, 2000; Scholes *et al.*, 2004). The detailed description for levels of achievement in the scoring rubric reduced uncertainty, hence, time was spent assessing the competence rather than trying to work out what the statements mean or what the levels are.

Time could have also been saved during documentation as the scoring rubric provided mentors with a clear idea of the objectives, allowing them to accurately identify areas for where they can direct the student within the scoring rubric rather than writing out long comments. This becomes useful when there is a need to write a developmental action plan; the scoring rubric is instrumental in providing the language to use in the documentation, enhancing mentors' assessment literacy and freeing them from spending time trying to develop their own interpretation to work out what to write in the action plan. This was illustrated in a comment made by a mentor that with the scoring rubric '*the interview was shorter*'. Reporting that the scoring rubric saved time for mentors is a substantial and novel finding in this study.

Conversely, some students reported that the scoring rubric was time-consuming. Arguably, this is due to the scoring rubric giving students ownership and responsibility for their own learning as discussed earlier. This assumption is justified by the fact that the students who claimed the scoring rubric was time-consuming, also stated that it allowed them to have *deeper* discussions and made it easier to

see what level they are at and how to improve their standards. Interestingly, one of the students recognised that the scoring rubric saved time for the mentor. This may suggest that the scoring rubric seem to shift the perceived responsibility of identifying and achieving learning outcomes in practice more towards the student, reducing the workload burden felt by mentors (Black, 2011; Hunt, 2016).

Additionally, the study revealed that lack of time is a problem for mentors rather than students, and that longer assessment is perceived by students to facilitate deeper learning and fairer assessment. This is a useful insight that could be explored further in future research.

7.2 Evaluating the overall usefulness of a scoring rubric

The aim of this study was to evaluate the extent to which a consensus-based scoring rubric was effective in making practice-based assessment of pre-registration nursing students fit for purpose. Ultimately, 'gatekeeping' entry to the nursing profession.

The literature about the effectiveness of scoring rubrics in assessment across many professional disciplines is well established (Jonsson and Svingby, 2007; Shipman *et al.*, 2012), including nursing (Wu *et al.*, 2015; Chong *et al.*, 2016; Llauro-Serra *et al.*, 2018). This current study is unique in designing and evaluating a scoring rubric to assess professional attitude, behaviour and responsibility competency statements. It specifically relates to the PLPAD but applies to nursing in general. The study established that a well-designed and consensus-based scoring rubric enhances the rigour of formative and summative assessment in this population. The study clearly demonstrates that the easily understood language describing the performance criteria raised the assessment literacy, hence, facilitated articulating well-defined criterion-based feedback provision, accurate documentation when forming developmental action plans, and justification of decisions about level of performance in the summative assessment. This is in contrast to decisions being based on intuition (Duffy, 2006; Black, 2011; Burden *et al.*, 2018).

These findings concur with Wu *et al.*'s (2015) study where a scoring rubric was used to assess simulated clinical practice, which was reported as offering mentors a means to provide ongoing feedback and give 'concrete' recommendations to guide decision-making. In contrast, Scholes and Albarran (2005) found that when assessors find the assessment tool vague, they respond by giving obscure feedback to conceal their own difficulty in understanding the assessment document, and subsequently, students cannot identify where their weaknesses are or how to amend them. Terry *et al.* (2017) reported similar findings stating that because the language was poorly understood, there was no shared sense of how to operationalise the competency standards, resulting in lack of or ambiguous feedback to students.

Duffy (2003) made a similar argument and reported that educational jargon makes it difficult for mentors to link aspects of performance with the learning outcomes to justify their decisions. Additionally, it can be argued that providing identical criteria in the scoring rubric for both the formative and summative assessments gave mentors repeated exposure to the criteria with the advantage of increasing their familiarity with what is being assessed at the summative point, a view shared by Robbins *et al.* (2018).

The enhanced rigour of assessment that the scoring rubric produced in this study could also relate to reducing mentors' biases or limiting their potential to apply hidden criteria not stated in the scoring rubric. Giving students access to the scoring rubric in advance allowed for shared understanding between the student and the mentor, hence, students were only held accountable for the stated criteria in the scoring rubric, limiting the opportunity to be assessed against unjustified criteria not stated in the scoring rubric. Thus, fairness of the assessment process to students and the improved rigour of practice-based assessment is a major finding of this study.

Although not tested statistically, the general perception of both mentors and students in this study was that the scoring rubric would improve both inter-rater

and intra-rater reliability. It is very likely that the increased reliability is due to the explicit language and well-defined performance levels. This would have eliminated discrepancies in interpreting the statements across the achievement levels and therefore assisted mentors to be consistent in their judgement irrespective of who assessed the performance. This echoes Jonsson and Svingb's (2007) finding that scoring rubrics assist raters in achieving high consistency when scoring performance tasks, since scoring rubrics can be seen as regulatory devices for scoring, therefore, scoring with a scoring rubric is more reliable than scoring without one.

While the complexity and fluid nature of clinical placements cannot be standardised and may be viewed as a threat to the consistency of assessment, this study reveals that by empowering students to be more responsible and confident regarding what is expected of them, they learn to take ownership of their learning and the power imbalance between the mentor and student begins to level out. This could refocus the mentor-student relationship towards a democratic and emancipatory critical dialogue based on mutual pursuit of understanding.

Subsequently, students would be more likely to practice deeper levels of reflection, and mentors would have more insight about students' perceptions and reasoning processes in unpredictable 'real life' clinical settings (Rennert-Ariev, 2005; Trede and Smith, 2012). In part, this current study shows that the scoring rubric allows both parties to work together to identify areas for further development, but more importantly gives mentors more insight to scrutinise the student's legitimacy to become a professional.

Within the context of this study, the success of the scoring rubric could also be attributed to the enhanced validity of practice-based assessment. Validity in educational research is often seen to involve evaluative judgement, therefore is not seen as a property of the test as such, but rather as an interpretation of the results (Messick, 1995). Moskal *et al.* (2000) explained that validity of scoring rubrics refers to the degree to which the evidence that supports the interpretations of the performance levels is accurate and appropriate, and added that there are three types of evidence commonly examined to support the validity of an assessment

instrument: content, construct, and criterion. Table 7.1 (overleaf) explains how the validity of the scoring rubric was scrutinised against Moskal *et al.*'s (2000) three measures.

Table 7.1: Validity of the scoring rubric against Moskal et al.'s (2000) criteria.		
Type	Description	How it was achieved in this study
Content	The extent to which a student's responses to a given assessment instrument reflects the student's knowledge of the content area that is of interest.	The scoring rubric was designed specifically for the professional attitude, behaviour and responsibility competency statements in the PLPAD. The iterative process used throughout the Delphi study ensured that all criteria were complete and reflected what the standards should mean.
Construct	An assessment instrument is complete and only measuring the intended construct.	Stakeholders' consensus was achieved, which means all attributes presented in the scoring rubric were relevant and precisely assess the related competence.
Criterion	The extent to which the students' performance on the given task may be generalised to other, more relevant activities.	Stakeholders' agreement was also obtained on the appropriateness of the level descriptors in distinguishing between different levels of competence that can be repeated in all clinical contexts.

Construct validity was further enhanced in this study due to the scoring rubric being well grounded in theory (Messick, 1995; Parratt *et al.*, 2016). The design of the scoring rubric was underpinned by the concepts of assessment *for* learning and authentic assessment.

Additionally, the rigorous methodological guidelines followed when conducting the Delphi to reach a consensus on the interpretation of the level descriptors used to develop the scoring rubric, enhanced face and content validity. Schuwirth and van der Vleuten (2011) emphasised the importance of considering validity as a subjective and qualitative judgement normally taken by experts in the domain being tested.

On the contrary, one might argue that mentors and students being briefed by the researcher on how the scoring rubric should be used could be viewed as providing training to use the tool, and therefore contributed to the positive outcomes. Such influence cannot be excluded since problems in knowing how tools should be used

are reported in the literature (Norman *et al.*, 2002). Nevertheless, Scholes *et al.* (2004) reported in their study of introducing portfolios, that even when a week was dedicated to prepare both students and mentors for using the tool, including detailed breakdown of the way to use learning contracts, and ways to understand the nature of the evidence to support learning, mentors and students still did not feel confident until they had actually used it in practice. Scholes *et al.* (2004) concluded that although the preparation was considered very helpful, confidence in using the tool only came from the experience of working with it in practice. Therefore, briefing participants on how to use the scoring rubric cannot fully explain its effectiveness.

7.3 Resonance with underpinning theories

The aim of this study was to make practice-based assessment of pre-registration nursing students fit for purpose through developing a scoring rubric and appraising its utility in enhancing interpretation between different levels of competence, facilitation of constructive feedback and crucially to examine if the scoring rubric strengthened the overall rigour of practice-based assessment. Ultimately, safeguarding the public from incompetent practitioners.

The theoretical frameworks that underpin this thesis and relate to the use of scoring rubrics (see section 3.4) are characterised by concepts that focus on the application of process and product. What this discussion chapter has established thus far is that using the scoring rubric to provide clear language and level descriptors to the attitude, behaviour and responsibility competency statements in the PLPAD was effective. This was evident in improving both the *process*, by emphasising the importance of formatively examining individuals' progress in relation to their starting point, and the *product*, to summatively judge if the performance met the required standards. The focus of the next section will be to evaluate whether the results in this study validate or further illuminate the theoretical frameworks underpinning the use of scoring rubrics in practice-based assessment.

Theories that emphasise process and product in practice-based assessment, such as Biggs's (2003) constructive alignment and Lave and Wenger's (1991) community of practice were considered relevant to the scoring rubrics' concept that assessment is an integral part of the learning cycle (see section 3.4). However, the distinguishable characteristics of scoring rubrics i.e. the clearly defined performance criteria written in easily understood language known in advance by students and mentors, resonate closely with the concept of assessment *for* learning which stresses the formative potential of the assessment through individualised feedback on performance (Black and Wiliam, 1998).

The transparency provided through making the scoring rubric available to both mentors and students in advance also resonates with the concept of authentic assessment by employing 'real world' tasks that mutually engage students and mentors in collaborative and meaningful activities (Wiggins, 1990). Therefore, the extent to which assessment *for* learning and authentic assessment theories explain the positive outcomes achieved in this study through using the scoring rubric will be appraised.

7.3.1 Assessment *for* learning

Changing views from assessment of learning to assessment *for* learning by emphasising the impact assessment has on learning has been long recognised in the literature (Boud, 1995; Black and William, 1998; Schuwirth and van der Vleuten, 2011), and requires adjusting the balance between formative and summative assessment (McDowell *et al.*, 2011). Participants' views expressed in this study are concurrent with the theory of assessment *for* learning by establishing that the scoring rubric strengthened the formative aspect of assessment, which enhanced learning, feedback and self-assessment.

Central to scoring rubrics being valuable in the formative assessment is their clarity and early introduction in the process. The explicit clarification of the learning outcomes in the scoring rubric made learning visible, reducing the commonly reported hindrance related to the ambiguity of the language used in competency

documents. Hence, as the concept of assessment *for* learning suggests, the scoring rubric helped mentors to understand and gauge where the student is along the continuum from where they were/are to where they need to be. It helped students see what they were aiming for and to understand what they needed to do to achieve those aims. Therefore, it helped them to take responsibility for their own learning and become more autonomous. This is consistent with the principal idea of assessment *for* learning through providing targets that focus on students reaching their full potential and the drive for deeper learning (Popham, 2001; Boud and Falchikov, 2006; Schuwirth and van der Vleuten, 2011; Haines *et al.*, 2013; Ali, 2013).

Comments from students and mentors reinforced the effectiveness of the scoring rubric in promoting learning by providing targets to exceed the required level instead of just passing. They indicated that enhancement of learning was not limited to how students can become competent, but also to understand accurately '*what constitutes excellence*', inspiring students who have high potential to excel. The comments expressed by students and mentors should make educators pause and deliberate whether the historic focus on identifying where the competent/not competent divider should be set has neglected investment in high achieving students. This question is worthy of a debate; however, this is outside the scope of this study and can be a suggestion for future research to consider.

Although the results of the scoring rubric evaluation correspond with the premise that assessment *for* learning enhanced learning, feedback and self-assessment, the validity of the theory in explaining the *role* of the scoring rubric in summative assessment is not clear. In other words, the theory clearly articulates how assessment *for* learning encourages progress, but the same cannot be said about its role in judging if students' progress is good enough to achieve the required outcome.

Conversely, it can be suggested that making the criteria identical and visible for both formative and summative assessments in the scoring rubric promoted the

formative assessment as a mechanism to prepare and help students to perform well in the summative assessment, but the theory does not sufficiently explain the accuracy of the summative decision. Schuwirth and van der Vleuten (2011) also acknowledged that this area of assessment *for learning* needs addressing.

7.3.2 Authentic assessment

As explored in the IR, the use of scoring rubrics is frequently associated with the concept of authentic assessment (Wiggins, 1989; Frentzos, 2013). Accordingly, the scoring rubric in this study was also underpinned by the authentic assessment concept, and the aim here is to examine if what this theory suggests concurs with, and explains, the findings of this study.

Authentic assessment involves assessing in the real world, focussing on the need for assessors to gain access to the context that underpins performance where problem-solving and critical thinking abilities are often used (Wiggins, 1989; Renner-Aeiev, 2005). Designing a scoring rubric in this study to authentically depict the real world of clinical practice was a fundamental stage. The ‘critically pragmatic’ decision to employ a Delphi method to achieve shared stakeholders’ interpretation of the competency statements, ensured they were contextualised to the realities of clinical practice, especially since mentors, students and service users actively participated in and agreed the competency statements’ interpretation.

This corresponds with Wiggins’s (1989) argument that what distinguishes authentic assessment is the designing of sophisticated criteria for judging performance, taking into account the actual performances the profession wants students to achieve. Therefore, it is essential to design criteria that replicate the actual challenges facing a person in the field. The intention of such “assessment of enablement” is to help students progress in handling complex tasks by learning to problem solve, argue critically and synthesise divergent viewpoints (Wiggins, 1989, p. 706).

Guided by the responses from mentors and students in this study, the scoring rubric allowed both parties to enter into a dialogue seeking mutual understanding that

invited exploration and reflection on their perspectives on practice. This concurs with the belief that, for the assessment to be authentic, it needs to reflect the intellectual work of practicing professionals characterised by active engagement, exploration, and inquiry on the part of the student, thereby enabling the assessor to watch the student pose, tackle, and solve slightly ambiguous problems (Wiggins, 1989).

Although competencies should conform to pre-determined standards set by professional bodies, current assessment documents are typically authoritative in dictating expectations (Rennert-Ariev, 2005). HEIs then interpret, design and produce their own documents to meet their own needs, subsequently using educational terminology unfamiliar to clinical practitioners (Rutkowski, 2007; Helminen *et al.*, 2016). In contrast, the scoring rubric designed in this study, although adhering to the requirements of the NMC standards, had statements written in positively worded, plain and non-threatening language. Additionally, the statements were constructed through a form of participatory social inquiry of those involved in pre-registration nursing education to reach common language for the professional attitude, behaviour and responsibility competency statements in the PLPAD.

The thread that has been running through in this chapter is the improved relationship between students and mentors, and how both parties used communication to establish competence. It is also clear that authentic assessment provides credible explanation that the authentic properties of the scoring rubric in this study allowed the student and mentor to work together in the reality and fluidity of the clinical settings and provided an insight into the student's ability to problem solve in the real world. Ultimately, the scoring rubric enabled transparency of mentors' judgements.

7.3.3 Alternative theoretical consideration - intersubjectivity

In recognition of the mentor-student relationship, that is strongly evident as a key factor throughout the study, this section explores whether an additional theoretical

framework, intersubjectivity, can be used to extend the debate of theoretical frameworks that may explain the use of scoring rubrics in practice-based assessment. The notion of intersubjectivity has been used in social science to refer to agreement between two or more individuals on a given set of meanings to provide a definition of the situation (Scheff, 2006) or to refer to the common-sense shared by people in their interactions with each other to interpret the meaning of elements of social and cultural life.

When examining dialogical theories of humans' sense-making, Linell (2009) identified that language is deeply intersubjective. He explains that when we talk, we perform cognitive and communicative actions in interacting with others and contexts, hence we always address our interlocutors, taking their perspective and orienting to what we think they think. Hence, in philosophy, the concept of intersubjectivity has its origins in the social theory of Jürgen Habermas. As discussed in section 4.2, Habermas (1984) assumes that language and meaning are the foundational component of the coordination of social action to achieve mutual understanding and '*truth*', using the expression 'intersubjectivity of mutual understanding' (Benjamin, 2013).

Within nursing, Cody (1995) argued that much of the discourse on objectivity and subjectivity could have been avoided had more attention been given to intersubjectivity. This view is echoed by Pierson's (1999) explanation that relationships within nursing cannot presume to rely solely on personal and intuitive perceptions, nor can relationships rely completely on objectively derived understandings:

"Within this relationship, teachers have the opportunity to know students as unique individuals - to be aware of their patterns of thinking, their usual actions and responses to client care situations. This knowledge and understanding, coupled with theoretical and technical knowledge, aids educators in guiding students through the process of teaching/learning related to nursing care." (Pierson, 1999, p. 299).

In this study, there has been a noticeable and significant thread where mentors and students reported that the scoring rubric gave them *shared purpose* that not only fostered growth and promoted learning and development, but also *stimulated* discussions and interactions to reach a *joint identification* of what constitutes 'competent' and build *mutual understanding* of that meaning.

Exploring the reasons for such persistent comments in this present study reveals the clear emphasis upon how both mentors and students perceived dialogue and mutual understanding as vital elements in practice-based assessment. Explaining this through the notion of intersubjectivity provides original understandings to emerge. In reference to the literature reviewed in chapter 2, as a consequence of the 'blurred' terminology of competency assessment, mentors have a subjective and implicit understanding of what a competent student looks like. The transparency of the scoring rubric made this implicit understanding more explicit, which in turn facilitated common understandings of competence. Understanding assessment in terms of intersubjectivity means that practice-based assessment is not considered a one-way process in which mentors observe performance. Assessment as an intersubjective act entails that assessment is thought of as a co-constructive process in which meaning is produced (Biesta, 1994).

The research by Black (2011) and Hunt (2014) establishes the overwhelming stress and anxiety mentors experience when, based on a relatively brief encounter with the student, they are required to judge (and predict) if that student will be a safe practitioner. Looking at this through the intersubjectivity perspective, a plausible explanation is that a likely source of the stress and anxiety for mentors is that they feel that making an accurate judgement about a particular student's legitimacy to become a professional is a huge burden and incredibly difficult to attain, especially when assessing attitudes and behaviours, as this assumes assessors can decipher what other people are thinking and feeling, i.e. "mind-reading". For example, is the student's belief or perception rational? To what extent does their behaviour explain what is really going through his or her mind?

With ambiguous language, mentors might avoid showing their lack of understanding, which can leave mentors having to make decisions mostly based on what they think (or hope) the student is thinking (meta perceptive - trying to think what the other person is thinking). Similarly, for the students, since they do not understand the terminology of competency statements, they would have no clear knowledge of what the mentors want to see in order to make a decision that the student is competent. Therefore, the student will also be trying to 'mind read' the mentors' perception about what constitutes an appropriate level of competence, so they can conform to it (meta-meta perspective - trying to think what the other person thinks he/she thinks).

In contrast, the statements plainly written in the scoring rubric provided the 'space' for dialogue and mentors had more insight into how the student is thinking rather than making assumptions or using their intuition. For the students, they reported having a clear idea about what is the acceptable level of competence at which they need to perform, and through dialogue with the mentor, they jointly agree the required level.

7.4 Chapter summary

This chapter integrated the quantitative and qualitative findings and, guided by the research aim and objectives, provided a critical examination of the effectiveness of a consensus-based scoring rubric in making practice-based assessment of pre-registration nursing students fit for purpose.

This chapter has justified the claim that the scoring rubric makes a positive contribution to practice-based assessment because it enhanced the clarity of the language and so aided mentors' and students' ability to recognise and distinguish between levels of competence. The greater accuracy and consistency of assessment provided by the scoring rubric led to increased rigour in the summative assessment of competence in the clinical settings. It also improved learning, self-assessment and feedback provision at the formative assessment stage. This was achieved through giving students positive reinforcement, whilst providing a clear

developmental framework to help them improve their performance through having clear targets towards which to work. Students could take ownership of their learning. In particular, the scoring rubric helped the more able students to achieve their desire to excel which, as the discussion in this chapter shows, the current system neglects.

The perceived or real unfairness that, sometimes, could lead to conflict between the mentor and student was reduced. The transparency of the scoring rubric facilitated mutual understanding over what needed achieving, using communication between mentor and student as a source of judgement.

The chapter also evaluated additional findings that were not directly explicit in the study aim and objectives. Most important of these, given the intrinsic nature of nursing which is relationship-based, the scoring rubric encouraged interaction and dialogue between them. This study has revealed, for the first time, that mentors worry about conflict developing in the mentor-student relationship as a result of the lack of clear understanding of language of competency statements in practice assessment documents.

Additionally, the scoring rubric was positively perceived as saving time, which, as existing literature shows, is one of the greatest challenges in clinical practice. Students liked the more valuable time that the scoring rubric created as they felt it was time spent on helping them to succeed or excel. Most importantly, as this discussion chapter has shown, the scoring rubric helps to address the practice-based assessment concerns raised by Duffy (2003) that most mentors fail to fail and Black (2011) that some mentors will fail students but struggle to evidence their decision.

The chapter evaluated the extent to which the theoretical frameworks of assessment *for* learning and authentic assessment were used to underpin this study and further our understanding of the findings. The concept of assessment *for* learning clearly explained enhancement of learning, self-assessment and feedback,

but does not sufficiently explain the accuracy of summative assessment. Authentic assessments tenets were found to be more compatible with the effectiveness of the scoring rubric in encouraging critical dialogue that is based on mutual understanding and produce more insightful judgements of students' performance in the real world.

Intersubjectivity theory was also found to provide a unique perspective on the mentor-student relationship and dynamics. The transparency of the scoring rubric made the implicit understanding more explicit, which facilitated common understanding of competence and provided the 'space' for dialogue, allowing mentors to have more insight into how the student is thinking rather than making assumptions.

The next chapter will be the concluding chapter and will provide direct answers to the research questions including the original contribution to knowledge. Recommendations for future practice and research will also be offered.

Chapter 8: Conclusion

8.1 Introduction

This final chapter summarises the study findings, how they address the original research aim and questions, and synthesises and confirms the new knowledge generated. The unique contributions to knowledge made by this study relating to the use of scoring rubrics in practice-based assessment are highlighted. The implications of this new knowledge will be valuable to HEIs providing pre-registration nurse education, placements providers, the Pan London Practice Learning Group that leads the PLPAD project, and the NMC as policy maker and regulatory body of nursing in the UK. The potential benefits to other HCPs globally will be discussed. The synthesis and recommendations take into consideration the newly published Standards for Student Supervision and Assessment (NMC 2018a) thereby reinforcing the enduring nature of the new knowledge. The chapter concludes by providing recommendations for further research.

8.2 Addressing the research aim

This study was initiated based on the concerns noted through the researcher's role as a lecturer teaching the NMC approved mentorship preparation programme. These concerns were regarding the lack of reliability and validity of mentors' assessment of pre-registration nursing students in practice placements, making practice-based assessment not fit for purpose, which ultimately could result in unsafe students joining the professional register, compromising public safety. A scoping exercise conducted locally with stakeholders identified several issues affecting the quality of the practice-based assessment process and outcome. Ambiguous terminology and difficulty in identifying performance levels in the practice assessment documents emerged as having significant impact (see section 2.3), and a critical review of the available international literature demonstrated that this was more than a local issue (see chapter 2).

Therefore, the way competency statements are interpreted merited further analysis and an integrative literature review was conducted which collated empirical and

theoretical literature to provide a broader and deeper insight into the challenges that mentors, students and nurse educators face in interpreting and assessing levels of competence. The IR establishes that central to the problem of practice-based assessment not being fit for purpose is a continuing inconsistency and difficulty in interpreting levels of performance when assessing nursing students in practice as well as difficulty providing constructive feedback for formative learning (Almalkawi *et al.*, 2018). Additionally, the review exposed that theories underpinning practice-based assessment require strengthening, and appropriate tools and taxonomies to assist mentors and students in the process are lacking (see section 3.3.4).

There is an urgent public safety need to develop a more robust method to improve practice-based assessment processes. The IR established a need for clearly understandable distinctions between different levels of competence and identified that a well-designed scoring rubric could be the solution (Almalkawi *et al.*, 2018). Other researchers have suggested conducting research to explore whether a scoring rubric, with a transparent and common language to interpret different levels of competence, offers a solution to the challenges faced in practice-based assessment (Jonsson and Svingby, 2007; Donaldson and Gray, 2012; Heaslip and Scammell, 2012; Shipman *et al.*, 2012; Frentsos, 2013). Therefore, underpinned by the concepts of assessment *for* learning and authentic assessment theories, the overall research aim in this study was to develop a scoring rubric based on stakeholders' consensus on interpretation of level descriptors for the professional attitude, behaviour and responsibility competency statements in the PLPAD, and evaluate the effectiveness of this consensus-based scoring rubric in making practice-based assessment fit for purpose. This is the first UK study with pre-registration nurse education to do so. Additionally, since assessing professional attitudes are notoriously challenging to define or measure (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016), this study is the first anywhere to design and evaluate a scoring rubric for professional attitudes and behaviours.

The study adopted critical pragmatism as an overarching ontological and epistemological approach. This approach continues to hold the pragmatic principles

of uncertainty about absolute '*truth*' and, inspired by Dewey (1966), postulates that ideas are used as instruments of inquiry that require active participation in processes to reconstruct events with alternative variations, followed by experimental testing of the *practical instrumentality* of the proposed change (see section 4.2).

Mixed research methods, comprising two phases, were implemented to achieve the study aim. The first phase used a Delphi method to reach stakeholder consensus on the interpretation of three level descriptors for the professional attitude, behaviour and responsibility competency statements in the PLPAD, which then formed the blueprint for the scoring rubric (see chapter 5). In the second phase, the developed consensus-based scoring rubric was evaluated in practice for its effectiveness in relation to three aspects: provision of clear language for the level descriptors, strengthening the rigour of competency assessment, and enhancement of learning, self-assessment and feedback provision (see chapter 6).

8.3 Answering the research questions

The primary research questions created in this study were to address the gap identified in the literature reviewed in the IR (see section 3.5), the answers to which constitute new and original knowledge. The research questions are directly answered in this section.

Question 1: Can consensus be achieved among stakeholders on how to interpret level descriptors for the professional attitude, behaviour and responsibility statements in the PLPAD?

Within the context of this study, the Delphi consensus method was selected to encapsulate the insights of different stakeholders and gain a collective interpretation of three level descriptors for the professional attitude, behaviour and responsibility statements within the PLPAD. This study was successful in reaching a very strong consensus (ranging from 86% to 100%) on shared interpretations of level descriptors for professional attitude and behaviour competency statements, which is a unique outcome of this study. These shared understandings informed the

construction of an original scoring rubric which was subsequently evaluated in practice for its effectiveness.

This study proves that the Delphi, as a consensus method, offers a highly structured and effective approach in facilitating democratic deliberation among participants with various roles who are geographically dispersed. Delphi is also a suitable method to determine if a consensus exists in interpreting professional attitude and behaviour competency statements, commonly reported in the literature to be abstract and open to interpretation, and therefore, notoriously challenging to define or measure (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016).

Question 2: Does a consensus-based scoring rubric improve interpretation of level descriptors?

This study confirms that the scoring rubric enhances understanding of the terminology of competency statements as well as the ability to recognise and distinguish performances at different levels. Both mentors and students perceived the scoring rubric to be effective in improving understanding and made implementation of the assessment process clearer. Therefore, the consensus-based scoring rubric is successful in addressing the ambiguity of terminology problems, which are reported frequently in assessment tools.

Having clear understanding of the language and level descriptors improves documentation and the evidencing of assessment decisions. Enhanced transparency, as found with the use of a scoring rubric in this study, facilitates mutual understanding over what students need to achieve. Mentors found the scoring rubric useful in providing a frame of reference that guided them to choose the right language, especially when there was a need to articulate an action plan and to communicate their decisions. Unlike current practice and existing assessment tools, which employ an achieved/not-achieved threshold, the grading process becomes more discriminating when mentors and students are provided with three level descriptors within a scoring rubric, allowing mentors to identify at

what level students are working. Additionally, the scoring rubric provides developmental scope for students to improve and excel. This latter point about more able students desiring to receive developmental guidance in order to excel is an original contribution to knowledge from this study and the scoring rubric is effective in providing that support.

Question 3: Does a consensus-based scoring rubric strengthen the rigour of mentors' assessment?

The scoring rubric assists mentors in making informed decisions about what is expected of students, as well as allowing them to justify their decisions. Assessment becomes more standardised and both mentors and students can consistently identify the required level of performance. The consensus-based scoring rubric enhances objectivity because the precise language and descriptors prevent decisions from being made on subjective assumptions or 'halo- horn' impressions (see section 2.3). This unique study shows that employing a well-designed scoring rubric gives mentors confidence that their judgements are fair and, a particularly encouraging finding, is students acknowledging that a scoring rubric helps them feel they have been treated fairly.

The scoring rubric also enhances the accuracy of practice-based assessment, unlike current assessment tools that, despite being based on taxonomies, provide only generic descriptors that lack specificity. The scoring rubric's success in enhancing accuracy is linked to its explicit description of performance levels written in an easily understood language, which formed objective criteria. Using identical criteria for both the formative and summative assessments gives mentors repeated exposure to the criteria with the advantage of increasing their familiarity with what is being assessed at the summative point. This study confirms that this increases the rigour of summative assessments. Additionally, giving students access to the scoring rubric in advance enables them to produce evidence that they have met the criteria, and restricts mentors from imposing unjustified criteria not stated in the scoring rubric. Thus, the improved fairness of the assessment process to students is a

recognisable benefit and another original contribution to knowledge from this study.

Question 4: Does a consensus-based scoring rubric enhance learning, self-assessment and feedback provision?

The scoring rubric is instrumental in providing structured and constructive feedback, especially when making action plans to help the ‘not yet competent’ student improve, and the more able student to develop beyond merely a ‘pass’. The provision of a structured and transparent developmental framework provides clear targets for both students and mentors. Having explicit criteria increases mentors’ knowledge and understanding of different levels of competence and enhances their assessment literacy, allowing them to explain and articulate what had been previously based on intuition. This study demonstrates that a well-designed consensus-based scoring rubric is perceived by students to improve learning and self-assessment.

The scoring rubric also provides encouragement through the positive language used during the formative stages of student learning. Unlike other forms of assessments that focus on highlighting incompetence, the scoring rubric allows personalised identification of areas for learning so guides students to become competent, steering them to reach their full potential, with some clearly striving for excellence as this study uniquely identifies.

The scoring rubric also facilitates students’ self-assessment. Transparent language and level descriptors as used in the scoring rubric allow students to gauge their own progress against the criteria and reflect on their own strengths and weaknesses. This helps them increase their autonomy, develop a better sense of ownership of their learning and be more pro-active in the assessment process.

To summarise, in answering the research questions, this current study establishes that a well-designed consensus-based scoring rubric enhances the rigour of formative and summative assessment of the professional attitude, behaviour and

responsibility competency statements in the PLPAD. A scoring rubric with easily understood language describing the performance criteria raises the assessment literacy, criterion-based feedback provision, self-assessment, accurate documentation, and justification of decisions about the level of performance in the summative assessment. This new knowledge has applicability beyond the local and national context and can be translated to practice-based assessment in other professions and countries.

8.4 Additional findings

In addition to answering the original research questions, several findings emerged from the study that were not directly explicit in the study objectives. These included issues related to mentor-student interaction and dialogue, and time.

8.4.1 Mentor-student interaction and dialogue

Central to nursing is the ability to foster positive relationships underpinned by effective communication. The scoring rubric gave mentors and students shared purpose that fostered growth and promoted learning and development. This original study provides an interesting dimension about utilising communication as a tool to establish competence rather than merely a competence to be achieved (see section 7.1.5). The scoring rubric stimulated discussions and interactions to reach a joint identification of what constitutes 'competent'. The importance of mutual understanding has not previously been given the attention it deserves in practice-based assessment, particularly, in light of the anticipated reduced time students will be spending with the practice assessor when the new NMC standards are introduced.

Having shared understanding facilitates opportunities for mentors and students to engage in critical thinking, allows mentors to make sense of students' *communicated* responses and rationales, and provides mentors with more concrete justification of students' level of competence when scrutinising their legitimacy to become a professional. This is in sharp contrast to current practices where decisions

about students' fitness for practice may be based on subjective perceptions and intuition (Duffy, 2006; Black, 2011). This study also revealed for the first time that mentors worry about conflict developing in the mentor-student relationship as a result of the lack of clear understanding of language used in practice assessment documents. Avoiding conflict is a recognised reason for failing to fail students of concern (Duffy, 2003; Gainsbury, 2010)

8.4.2 Time

This study found that the scoring rubric saves time. This is a very important finding carrying potential to have a significant impact on practice-based assessment since the scoping exercise (see section 2.3) and a vast amount of international literature identify lack of time as a key and chronic barrier to effective assessment in practice. The scoring rubric's hallmark of transparency and clarity of the language and level descriptors reduces uncertainty, subsequently time can be spent assessing competence rather than, as others (Neary, 2000; Scholes *et al.*, 2004) have reported, trying to deconstruct competency statements to make sense of what they mean or what the levels are.

Time is also saved during documentation (see section 7.1.6). When mentors have clearly articulated and agreed descriptors of student performance, they can employ the language from those descriptor statements to provide succinct feedback and evidence their assessment decisions. Having clear distinction between levels of competence can free mentors from spending time trying to work out how to express what is required in a student's developmental action plan.

For students, the scoring rubric seemed to be time-consuming. This should be viewed as a positive outcome since they considered the extra time to be valuable in helping them to succeed or excel. Students also identified that the scoring rubric gave them more ownership and responsibility for their own learning and this is a valuable contribution to knowledge since previous literature has indicated that students may be too passive in their engagement with learning (Brown *et al.*, 2005; Levett-Jones *et al.*, 2009). This study is unique in revealing that lack of time is a

problem for mentors rather than students, and that students perceive longer assessments as facilitating deeper learning.

8.5 Strengths and limitations

This study used mixed methods methodology consisting of two phases where quantitative and qualitative data were collected and analysed. All findings were integrated and synthesised to reach a greater understanding of the usefulness of the scoring rubric in enhancing practice-based assessment.

One of the strengths of this study is the robust philosophical and theoretical underpinning. This study has the advantage of making a contribution to nursing research by using a critical pragmatist's perspective. It has provided evidence that a critical pragmatist approach is valuable by enabling the development of a scoring rubric that ensured social involvement through a two-round Delphi method in deciding what is *truth* and *knowledge* when interpreting the professional attitude, behaviour and responsibility statements in the PLPAD. This was followed by evaluating the *practical instrumentality* and usefulness of the scoring rubric in clinical practice when used by mentors and students.

Another strong point in this study is related to the solid theoretical framework which unpinning the use of the scoring rubric in practice-based assessment. The concepts of assessment *for* learning, authentic assessment and intersubjectivity have provided greater understanding of the processes underlining the positive outcomes achieved in this study when the scoring rubric was evaluated in practice.

Methodologically, the mixed methods approach added rigour by employing established and relevant processes for both the quantitative and qualitative methods in both phases, drawing on the strength of both approaches to stimulate an inquiry that complemented one paradigm with another. However, there were some limitations in using the Delphi method in this study. The method has been criticised by believers of the scientific approach mainly in areas such as design and administration (Linstone and Turoff, 1975; Keeney *et al.*, 2011). Their concerns were

partly reflected in this study, as little clear guidance could be found in published literature on methodological standards. To mitigate for this, every effort was made to maintain a rigorous process by not modifying the classic Delphi method through mainly using a qualitative first round. Other strategies used included piloting the questionnaires, using a representative sample, keeping the process short and defining clear consensus criteria.

Although this study drew on a national sample of stakeholders, the generalisability of the findings may be affected by recruiting student nurses from one university; most of the panel members were based in London (using the PLPAD), which was necessary as the PLPAD was used in the London region at the time of conducting the study. Therefore, the findings have some limitations on generalisability beyond the single London HEI and further research with a wider nationally-representative sample is needed. While the purposive sampling method used in both phases was necessary to reach the target population, it is also considered a limitation of this study due to potential to introduce bias in the sampling process.

The uncertainty of the response rate in the Delphi phase is a potential limitation in this study. Accurately calculating the response rate was complicated by taking on board a recommendation made by a university reviewer advising to make the email invitation a one-stage process by including the information sheet and the link to the questionnaire with the invitation email. Although this seemed to be a good idea at the time, a shortcoming was recognised later when it was not possible to differentiate between those who agreed to participate and those who chose not to. However, there were sufficient participants as the proposed number was met, and more importantly, a good representation of stakeholders was reached. Additionally, regarding the potential inclusion in the sample of mentors who find failing students difficult, it is a possibility that their interpretations could adversely affected the outcome of the consensus. This potential limitation was mitigated by the very nature of the iterative process of the Delphi where participants refined or reconsidered their opinions from round to round based on the opinions of other panel members.

In the first round of Delphi, despite piloting the questionnaire to ensure clarity, the design led to some respondents misunderstanding a few questions, especially the negatively worded ones. However, this had a negligible effect on the consensus reached and was taken into consideration when designing the questionnaire for the second phase of the study. Furthermore, in the evaluation phase, there were some participants who strongly disagreed that the scoring rubric was useful, but they did not provide comments in the free-text box provided in the questionnaire to explain their reasons. Therefore, not knowing whether respondents did not find the scoring rubric useful or they misinterpreted the direction of the Likert scale is a limitation.

A validation study using psychometric testing of the scoring rubric was one of the original objectives of this research to test construct validity using principal component analysis, with Cronbach's alpha analysis to test internal consistency. However, this was found to be beyond the scope of this research study, following advice from a professor of statistics who suggested that conducting psychometric testing would be more appropriate after repeated experiments of the scoring rubric as a post-doctoral project.

8.6 Conclusions arising from the study

The literature to date highlights that ambiguity in understanding and interpreting the language in practice assessment documents or the ability to discriminate between performance levels is a significant barrier to assessing students' legitimacy to join the professional register, making practice-based assessment not fit for purpose. This original study targeted professional attitude and behaviour competency statements, often reported to be notoriously difficult to define and measure (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016), and took a critical pragmatist approach in generating consensus on a collaborative interpretation of level descriptors for competency statements. These statements were subsequently incorporated into a scoring rubric and evaluated for their effectiveness in practice. Therefore, this study is unique in involving mentors and students in interpreting and evaluating competency statements.

The findings show that, in the context of this study, the consensus-based scoring rubric is effective in making practice-based assessment fit for purpose. The scoring rubric provides clear language and level descriptors for the competency statements, enabling both mentors and students to mutually understand what is required to be achieved and how it can be achieved. This improves students' self-assessment and the provision of accurate developmental feedback from their mentors. Most importantly, this unique scoring rubric for professional attitudes and behaviours enhances the rigour and fairness of assessment. Additionally, this study reveals other advantages of using a well-designed scoring rubric extend to include empowering students, encouraging deeper mentor-student interaction and dialogue, and saving mentors' time. However, some of these conclusions are tentative and will need strengthening by further national studies with a wider group of stakeholders before confirming the impact of the scoring rubric in assessing pre-registration nursing students.

This thesis provides a theoretical insight into practice-based assessment. The concepts of assessment *for* learning and authentic assessment broaden the understanding of practice-based assessment in the complex and dynamic environment of clinical practice. However, the participants' frequent reference to improved mutual understanding and dialogue is a clear indication that current assessment practices are mostly a one-way process, with mentors struggling to decipher whether a student is competent or not and finding this an overwhelming and stressful experience.

Examining mutual understanding in practice-based assessment through the intersubjectivity perspective, allows deeper insight into how the scoring rubric was instrumental in enhancing mentor-student mutual understanding and dialogue. The dialogue that is created by a well-designed, transparent scoring rubric allows mentors to make sense of students' patterns of thinking and responses to the demands and dilemmas of continuously changing real-life situations. This allows

beliefs and values to be more exposed, hence, provides mentors with better insight when making judgements of student's fitness for practice.

8.7 Contributions to knowledge

The new knowledge that has been generated in this study makes several contributions to the field of practice-based assessment of student nurses and nursing theory, and it is highly likely that the findings are generalisable to other groups assessed in practice. This includes contributions to the knowledge base in relation to the use of scoring rubrics in assessing professional attitude, behaviour and responsibility competency statements; the application of Delphi consensus method to generate a consensus on level descriptors for professional values; providing a theoretical understanding of practice-based assessment through assessment *for* learning and authentic assessment lens, and being guided by the ontological and epistemological tenets of critical pragmatism.

8.7.1 Contributions to the field of practice-based assessment

A substantial contribution to the field of practice-based assessment is directly in response to the literature frequently reporting that the terminology used in practice assessment documents is vague, packed with academic jargon, and differentiating between levels of performance continued to be subjective (Almalkawi *et al.*, 2018). This study is unique in providing evidence that clear and unambiguous language to describe competency statements that distinguish between different levels of competence is achievable with the use of a well-designed consensus-based scoring rubric.

Although there is evidence that scoring rubrics are used in nursing, this study is the first to design a scoring rubric specifically to enhance accurate assessment of professional attitudes and behaviours. There is a public safety need for practice-based assessment to be fit for purpose not only within nursing but also across other health care professions. This is an important contribution to knowledge since the literature clearly reports that making decisions about attitudes and behaviours is

notoriously challenging to define or measure (Fitzgerald *et al.*, 2010; Miller, 2010; Butler *et al.*, 2011; Hunt, 2014; Strauss, 2016).

The ultimate contribution that this study makes is the impact of the scoring rubric on making practice-based assessment fit for purpose, as evidenced by mentors and students reporting that assessment of competences was objective, rigorous and fair. Reports by mentors that the scoring rubric assisted them in making informed decisions and justifying them, is an important contribution since national and international research continues to indicate that assessment of students in practice is based on intuition, thus, mentors struggle to justify that their decisions are valid. Additionally, this study is unique in illustrating that the scoring rubric provides a frame of reference for students, facilitating self-assessment, giving them autonomy, a sense of ownership and insight into their strengths and weaknesses. They were able to see the fairness of the decisions made and felt liberated to challenge decisions they considered deviated from the transparently stated criteria.

This study stresses the importance of formative assessments in providing individualised feedback on performance which is an area where there has been a paucity of literature. This study is exclusive on two fronts. First, by emphasising the use of positive language to show students how to become competent rather than highlighting incompetence. Second, by offering encouraging and personalised information to steer well-performing students to reach full potential and strive for excellence.

This study also demonstrates that the transparent criteria in the scoring rubric were instrumental in saving mentors time when making action plans to manage poor performance, and enhancing their assessment literacy when documenting evidence. The scoring rubric allowed mentors to accurately identify areas for development by referring to the scoring rubric, thus freeing them from spending time trying to work out what to write in the action plan. Reporting that the scoring rubric saved time for mentors is a novel finding in this study. Additionally, the study reveals that lack of time is a problem for mentors rather than students, and that longer assessments

are perceived by students to facilitate deeper learning and fairer assessment. Students' perceptions that the scoring rubric made the assessment fairer is a recognised benefit in this study.

8.7.2 Contributions to the field of nursing philosophy and research

This research study contributes to nursing philosophy by being the first to adopt the contemporary critical pragmatist philosophy. Critical pragmatism tenets provided a unique perspective and, as explained in chapter 4, were found to be the most suited overarching ontological and epistemological approach. Critical pragmatism aligned with the approach taken to design the scoring rubric by employing an alternative variant that ensured social participation in deciding what is *truth* and *knowledge* when interpreting the statements in the PLPAD. This was followed by evaluating the *practical instrumentality* of the proposed change experimentally as guided by the philosophy.

This study is also the first to use a Delphi consensus method to elicit stakeholders' interpretation for the professional attitude, behaviour and responsibility statements in the PLPAD. Although it has been used in nursing, to date, Delphi consensus methods have not been utilised in pre-registration nursing education or any other HCPs to inform professional attitudes and behaviours descriptors of a scoring rubric. The classic e-Delphi method used in this study is fully described within this thesis. Others could replicate the method to develop their own scoring rubrics.

8.7.3 Contributions to the field of nursing theory

Whilst being underpinned by assessment *for* learning and authentic assessment theories, this research study contributes to nursing theory by being the first to explore the applicability of the theory of intersubjectivity to further the understanding of practice-based assessments. Mentors and students frequently commented that the scoring rubric improved shared understanding, emphasising how both mentors and students perceived dialogue and mutual understanding as a vital element in reducing conflict and misunderstanding in practice-based assessment. This study identified that the anxiety experienced by mentors when

assessing students' competence in practice stems from insufficient meaningful interaction coupled with a lack of clear and transparent criteria. Mentors described the process of judging student's legitimacy to become a registered professional as overwhelming, trying to decipher if the student thinking processes are rational, and to what extent their behaviour explains what is really going through their minds (meta-perspective). Similarly, students struggle to work out their mentors' perception about what constitutes appropriate levels of competence. Hence they try to 'mind read' their mentors' perception so they can conform to it (meta-meta perspective).

To date, this is the first study to report that the clear and transparent criteria in the scoring rubric stimulates mentors and students' interactions, encouraging them to engage in critical thinking and mutually agree what constitutes 'competent'. The scoring rubric allows mentors to make sense of students' *communicated* responses and rationale and form a more concrete justification of students' level of competence and their safety to the public. Therefore, explaining dialogue and mutual understanding through the lens of intersubjectivity provides an original insight to emerge, which can be explored further in future studies.

8.8 Recommendations

It is important to note that the recommendations based on the outcomes of this study are relevant globally to other professions that incorporate assessing competencies in practice. The recommendations are aimed at any education institution that delivers pre-registration nurse education and mentorship preparation programmes as well as healthcare placement providers. They are also relevant to the Pan London Practice Learning Group responsible for designing the PLPAD and the NMC as the regulatory body for pre-registration education and mentor preparation. The recommendations take into consideration the new NMC standards for student supervision and assessment published in May 2018 to be implemented in 2019. Recommendations for future research are also suggested.

8.8.1 Recommendations for educational institutions

The results of this research show that interpretation of competency statements and performance levels can be enhanced by promoting shared understanding and dialogue, clarifying what the statements mean and what constitutes an appropriate level of competence. It is therefore recommended that education institutions incorporate 'interpreting competency statements' into the pre-registration nursing curriculum and mentorship preparation programmes. This could be through workshops that involve students, mentors and academics to participate in reaching shared understanding of how the competency statements and performance levels should be defined in the practice assessment documents.

This study provides evidence that designing competency statements with clear language and performance levels to replicate real life, although time consuming, enhances practice-based assessment. Therefore, a strong recommendation for education institutions is to consider adopting scoring rubrics in their curriculum across all years of their programmes. This recommendation is supported by comments from the students and mentors who participated in this study, which provided a strong indication that the current system may not be fit for purpose and needs to change.

A Delphi process that facilitates reaching a consensus is highly recommended in order to obtain agreement on interpretation of all the competencies for the pre-registration programmes, as it was found effective in reaching stakeholders' agreement in this study. Involving mentors and students in the process should be central. It is understandable that this will be challenging and complex but should be achievable since the scoring rubric in this study was developed by one part-time researcher with limited resources.

Recommendations also extend to mentor preparation programmes to incorporate workshops and training on how to make the most of using scoring rubrics. This should include learning how to decide the appropriate level, provide constructive feedback, appropriate documentations and crucially how to encourage dialogue

with students to reach deeper and more critical insights. The newly published NMC Standards for Student Supervision and Assessment (NMC, 2018a) abandon the dual role where mentors are simultaneously the facilitator and the assessor. These two roles will be separated into a practice supervisor and a practice assessor who cannot be the same person. Additionally, the current mentorship preparation programmes are no longer a requirement. Nevertheless, the practice assessor will continue to require preparation to achieve outcomes set in the new NMC standards. Despite healthcare providers still being unsure about how practice assessors will be prepared to meet the new NMC outcomes, the recommendations presented here are transferable and applicable to the new assessor role.

8.8.2 Recommendations for healthcare placement providers

The literature acknowledges that lack of time is a significant barrier to effective mentoring, which led the DH (1999), the NMC (2008a, 2018e) and the RCN (2015) to call for protected time (see chapter 2). This study revealed two important issues that are likely to have implications for practice. First, this current study is novel in establishing that a scoring rubric is a successful strategy to save time for mentors by reducing uncertainty and enhance assessment literacy. Second, lack of time is a problem for mentors rather than students, and that longer time is perceived by students to facilitate deeper learning and fairer assessment.

Therefore, a recommendation for healthcare placements providers is that they should fundamentally consider introducing new ways of working that challenge the norm of students relying on their mentors to facilitate the learning, and the perception that mentors need to be freed to 'teach' students. Mentors should be trained in practicing appropriate ways of using the scoring rubrics to guide students' learning, shifting the responsibility more towards the student. Students will also need to be trained to utilise the scoring rubric to self-assess their own learning needs and to identify the required learning outcomes in practice. Guided by the evidence identified in this study, both mentors and students can be expected to benefit from this new way of working; mentors' time will be saved, and their workload reduced, and at the same time, students will gain deeper learning and

fairer assessment. However, the ultimate benefit has to be facilitating mentors' and students' engagement in critical thinking, allowing mentors to make sense of students' communicated responses and rationale, and therefore, when scrutinising their legitimacy to become a professional, mentors would have more concrete justification for their decisions.

To enable this, it is recommended that healthcare placement providers consider assigning scoring rubrics champions, who can lead on promoting scoring rubrics and train mentors to employ them appropriately to achieve the desired learning outcomes and accurate assessment. Preparing practice assessors for their new role provides a good opportunity to introduce this change. Establishing a forum for peer support and sharing of good practice is also a suggestion.

8.8.3 Recommendations for the NMC and the Pan London Practice Learning Group.

The new NMC Standards for Students Supervision and Assessment (NMC, 2018a) cover learning and assessment in practice and propose radical and ambitious changes to the way HEIs and their practice partners educate and assess pre-registration nursing students. This includes introducing the roles of practice supervisor, practice assessor and academic assessor to work in a tripartite model to assess and confirm students' practice and academic achievement. Currently, the Pan London Practice Learning Group along with stakeholders from across London are working together to develop and agree a unified approach to the implementation of these roles.

However, the new practice assessor role will continue to be performed by registered nurses who will "conduct assessments to confirm student achievement of proficiencies and programme outcomes" (NMC, 2018a, p. 9), with less emphasis on preparing them for the assessor role in comparison to the previous NMC standards (NMC, 2008a). Therefore, it is difficult to envisage how the new standards will make a real change to the current problems facing practice-based assessment fitness for purpose, specifically the practice assessors' ability to interpret

competency statements or discriminate between performance levels. Subsequently their ability to safeguard the public from incompetent students joining the register remains questionable. Acknowledging that those in the new assessor roles are likely to continue to experience difficulties in interpreting competency statements and performance levels, scoring rubrics are recommended to be adopted when implementing the new standards.

Furthermore, the introduction of the tripartite model in the new NMC standards is likely to be problematic in view of the reported inconsistencies in the way students, mentors and university lecturers interpret competency statements (Dolan, 2003). Participants in this study illustrated that using the scoring rubric enhances consistency, making the assessment more standardised between mentors, as well as providing uniformity of a shared mentor-student understanding. Therefore, the scoring rubric is further recommended and would be a potentially useful instrument to improve consistency within the tripartite scheme when assessing students in practice placements.

It is also worth noting that the role of the practice assessor in the new NMC standards does not specify time that should be spent with students, but it is projected to be significantly shorter than the time mentors currently spend with their student since the dual role will be discontinued (NMC, 2018a). The implications for this are yet to materialise, however, since the scoring rubric saved time for mentors in this current study, this provides further support for recommending scoring rubrics when implementing the new NMC standards.

The exceptional work of the Pan London Practice Learning Group in developing a single document for use in the London region is acknowledged and commended. It is encouraging that the document is gaining interest from other HEIs outside London who expressed their interest in joining in. Such success brings the vision of developing a national practice assessment document a step closer to becoming a reality, which has been long called for in the literature (Norman *et al.*, 2002; O'Connor *et al.*, 2009; Cassidy *et al.*, 2012). It was also recommended by the Health

Education England's (HEE) project as a strategy to reduce pre-registration attrition and improve retention rates (HEE, 2018). However, what the present study found is that the 'achieved/not achieved' classification currently adopted in the PLPAD diminishes opportunities for using the assessment process as a source of feedback and learning.

The Pan London Practice Learning Group are working with HEIs and practice placement providers on a new PLPAD based on the new NMC Standards for pre-registered nursing education (NMC 2018b), underpinned by the values of The Code (NMC, 2015, 2018d). The group recently published the first draft, which continues to use the achieved/not achieved classification, despite receiving feedback from stakeholders that the language in the current PLPAD is complicated (Fish, 2018). Therefore, a recommendation arising from this study is to develop scoring rubrics for all the competency statements and across all years in the new PLPAD, which should include stakeholders in interpreting three level descriptors for each competency statement using positive language. This will ensure mentors and students can understand the language and will provide opportunities to identify areas for learning and plan how to address them, as well as offering well-performing students a developmental range with scope to excel.

8.8.4 Recommendations for future research

This study shows that within one HEI a scoring rubric is effective in providing clear language and level descriptors for the competency statements, which enhances the rigour and fairness of assessment, and improves provision of accurate developmental feedback and self-assessment. Further testing of the results of this study with a wider nationally representative sample is encouraged, as this would confirm the findings of this study, on the premise that retesting would make these findings more established. This could include multi-centre studies that apply the scoring rubric in a variety of clinical settings and with different professional groups. To gain more insight, it is recommended that future researchers emphasise to participants, especially if they strongly liked or strongly disliked the scoring rubric, the importance of giving reasons in the comments box. Additionally, participants

could be offered the opportunity to take part in a focus group to discuss their responses, positive or negative, in more depth.

Validation studies to perform psychometric testing of the scoring rubric is another recommendation. These commonly involve testing construct validity using principal component analysis, and Cronbach's alpha analysis to test internal consistency. The introduction of the practice supervisor, practice assessor and academic assessor roles in the new NMC standards (NMC, 2018a) will necessitate establishing consistency of the assessments conducted by different people, therefore the validation study should also examine the scoring rubric's ability to enhance inter-rater reliability within the tripartite scheme.

The findings in this study raise questions whether the previous focus on identifying where the competent/not competent divider should be set has neglected investing in high-achieving students. This question is worthy of future research to study high-achieving student experiences and feelings within the current practices that focus mainly on competence levels. Furthermore, investigating whether feelings of 'not getting a fair deal' in assessment contribute to dissatisfaction and attrition warrants further research.

Additionally, this study revealed that the lack of time, frequently reported as a major barrier in practice-based assessment, is a problem for mentors rather than students, and that longer time is perceived by students to facilitate deeper learning and fairer assessment. This is an important finding that could be explored further in future research to examine the importance of time for students' learning and development.

The contextual literature review in this study identified that to date there were no studies looking at the relationship between failing to fail students and the increasing referral rates to the NMC (see section 2.2.2). Researching this area is recommended to establishing whether such a relationship exists.

8.9 Dissemination

The new knowledge generated in this study about the effectiveness and usefulness of scoring rubrics in making practice-based assessment more fit for purpose will be presented to the relevant stakeholders. The researcher has already started engaging in discussions with the Pan London Practice Learning Group to present the findings of this study. As discussed earlier, the group has launched the new draft PLPAD document in light of the new NMC standards (2018a, 2018b), therefore, this provides a good opportunity to share the usefulness of using scoring rubrics and suggest developing scoring rubrics for all the competencies in the new document.

Dissemination of the thesis will also be through presenting at conferences and publication in relevant international nursing journals, specifically with education and research focus. The IR has already been published in *Nurse Education Today* (Almalkawi *et al.*, 2018). Using a Delphi methodology to develop a scoring rubric for professional values is unique and this method has been presented at the RCN education conference (Almalkawi *et al.*, 2017). A draft manuscript is currently being written for the *Journal of Advanced Nursing*. Employing critical pragmatist philosophy in nursing is an exclusive contribution and it is proposed that this will be disseminated through publication in *Nursing Research Journal*. A summary of the whole study will be disseminated through publication in a nursing journal with an international audience to encourage global cooperative work on the use of scoring rubrics in practice-based assessment.

8.10 Thesis conclusion

In nursing, as with other HCPs globally, assessing competence in practice is known to be problematic and not fit for purpose. This may result in practitioners joining the professional register and posing a risk to the public. Factors contributing to poor assessment in practice are complex, however, opaque terminology and inability to identify what constitutes appropriate levels of competence especially related to assessing attitudes and behaviours have a major impact on the quality of practice-based assessment.

This study is the first to uncover that what actually concerns mentors when the language used to describe competencies is unclear, is that the lack of clarity is a source of 'conflict and misunderstanding.' The lack of clear and transparent criteria when assessing students' competence in practice raises mentor anxiety and reduces meaningful interaction with their student. This study identifies that mentors found the process of judging student's legitimacy to become a professional overwhelming as they struggled to decipher whether the student's thinking processes were rational and to what extent their behaviour explained what they were thinking. From a student perspective, this study found that they struggle to work out what mentors consider constitutes appropriate levels of competence and try to 'mind read' their mentor to conform to expectations.

The lack of clarity is unnecessary since this study establishes that stakeholders can reach consensus on interpretations of level descriptors for competencies related to attitudes and behaviours to produce a scoring rubric. Testing of the consensus-based scoring rubric showed that it enhances the quality of practice-based assessment. Assessments that are fit for purpose can achieve the goal of ensuring better public protection. Mentors reported that the scoring rubric assisted them in making informed judgements and they were able to justify their decisions. Students found the scoring rubric facilitated self-assessment, gave them autonomy, a sense of ownership and insight into their strengths and weaknesses. They were able to see the fairness of the decisions made and felt liberated to challenge decisions they considered deviated from the criteria. Therefore, the ultimate contribution that this study makes is to demonstrate that fit for purpose practice-based assessment is achievable by using a consensus-based, transparent scoring rubric with clear descriptors that support formative learning as well as summative assessment and will be supported by mentors and students because it is objective, rigorous and fair.

References

- Addams, J. (1902) *Democracy and Social Ethics*. New York: Macmillan.
- Addams, J. (1910) *Twenty Years at Hull House, with Autobiographical Notes*. New York: Macmillan.
- Aggleton, P., Allen, M. and Montgomery, S. (1987) Developing a system for the continuous assessment of practical nursing skills. *Nurse Education Today*, 7(4), pp. 158-164.
- Akins, R.B., Tolson, H. and Cole, B.R. (2005) Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC medical research methodology*, 5(1), p. 1.
- Ali, J.M. (2013) Getting lost in translation? Workplace based assessments in surgical training. *The surgeon*, 11(5), pp. 286-289.
- Allan, H. (2010) Mentoring overseas nurses: barriers to effective and non-discriminatory mentoring practices. *Nursing ethics*, 17(5), pp. 603-613.
- Allen, D. and Tanner, K. (2006) Rubrics: tools for making learning goals and evaluation criteria explicit for both teachers and learners. *The American society for cell biology*, 5(3), pp. 197-203.
- Almalkawi I., Jester R., and Terry L. (2017) Developing a rubric to enhance mentors' assessment of student nurses' practice performance: A Delphi Process. RCN Education Forum International Conference and Exhibition. Nursing education and professional development: the global perspective, Cardiff, 21-22 March.
- Almalkawi, I., Jester, R. and Terry, L. (2018) Exploring mentors' interpretation of terminology and levels of competence when assessing nursing students: An integrative review. *Nurse education today*, 69(October), pp. 95-103
- Ashworth, P.D., Gerrish, K., Hargreaves, J. and McManus, M. (1999) 'Levels' of attainment in nursing practice: reality or illusion? *Journal of Advanced Nursing*, 30(1), pp. 159-168.
- Aston, L. ed., (2010) *The student nurse guide to decision making in practice*. UK: McGraw-Hill Education.
- Avery, A., Savelyich, B., Sheikh, A., Cantrill, J., Morris, C., Fernando, B., Bainbridge, M., Horsfield, P. and Teasdale, S. (2005) Identifying and establishing consensus on the most important safety features of GP computer systems: e-Delphi study. *Informatics in Primary Care*, 13(1), pp. 3-11.
- Axley, L. (2008) Competency: a concept analysis, *Nursing forum*, 43(4), pp. 214-222.
- Baillie, L. (1993) Factors affecting student nurses' learning in community placements: a phenomenological study. *Journal of Advanced Nursing*, 18(7), pp. 1043-1053.
- Baillie, L., Fish, J., Barclay, J., Case, D., Crussell, J., O'Brien, M., Icheku, V., Humphries, P., Thompson, S. and Mitchell, J. (2016) Assessing nursing students in practice: a mixed method evaluation of a unified assessment document. Poster presentation. Available from: https://www.heacademy.ac.uk/system/files/downloads/baillie-assessing_nursing_students.pdf [Accessed 21st May 2017].
- Barnett, R. (1994) *The limits of competence: knowledge, higher education and society*. Open Bristo: University Press.
- Basit, T. (2003) Manual or electronic? The role of coding in qualitative data analysis. *Educational research*, 45(2), pp. 143-154.

- Beasley, C. (2009) High Impact Actions for Nursing and Midwifery. The Essential Collection <https://www.qualitasconsortium.com/index.cfm/reference-material/innovation/high-impact-actions-for-nursing-midwifery/> [Accessed 12th May 2016].
- Benjamin, J. (2013) *The bonds of love: Psychoanalysis, feminism, & the problem of domination*. New York: Pantheon.
- Benner, P. (1984) *From novice to expert*. CA: Menlo Park.
- Benner, P. and Tanner, C. (1987) Clinical judgement: how expert nurses use intuition *American Journal of Nursing*, 87(12), pp. 23–31.
- Bettany-Saltikov, J. (2012) *How to do a systematic literature review in nursing: a step-by-step guide*. UK: McGraw-Hill Education.
- Biesta, G.J. (1994) Education as practical intersubjectivity: towards a critical - pragmatic understanding of education. *Educational Theory*, 44(3), pp. 299-317.
- Biggs, J. (2003) Aligning teaching and assessing to course objectives, *Teaching and Learning in Higher Education: New Trends and Innovations*, 2(April), pp. 13-17.
- Bindal, T., Wall, D. and Goodyear, H.M. (2011) Trainee doctors views on workplace-based assessments: are they just a tick box exercise? *Medical teacher*, 33(11), pp. 919-27.
- Bishop, G (1987) Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly*, 51(2), pp. 220-223.
- Black, S. (2011) Being a Mentor Who Fails a Pre-registration Nursing Student in their Final Placement: Understanding Failure. PhD Thesis. London, London South Bank University.
- Black, P. and Wiliam, D. (2006) *Developing a theory of formative assessment*, in: J. Gardner (Ed.) *Assessment and learning*. London: Sage.
- Black, P. and Wiliam, D. (1998) *Inside the Black Box: Raising Standards through Classroom Assessment*. London: School of Education, King's College.
- Bloom, B.S. (1956) *Taxonomy of educational objectives*. Vol. 1: Cognitive domain. New York: McKay, pp.20-24.
- Boerner, K.E., Coulombe, J.A. and Corkum, P. (2015) Core competencies for health professionals' training in pediatric behavioral sleep care: a Delphi study. *Behavioral sleep medicine*, 13(4), pp. 265-284.
- Bondy, K. (1983) Criterion-Referenced definitions for rating scales in Clinical Education, *Journal of Nursing Education*, 22(9), pp. 376-382.
- Bondy, K. (1984) Clinical evaluation of student performance: the effects of criteria on accuracy and reliability. *Research in nursing and health*, 7(1), pp.25-33.
- Bork, C.E. (1993) *Research in Physical Therapy*. Philladelphia: J.B. Lippincott Co
- Boud, D. (1995) Assessment and learning: contradictory or complementary, in: P. Knight (eds.) *Assessment for Learning in Higher Education*. London: Kogan Page, pp. 35-48.
- Bowling, A. (2014) *Research methods in health: investigating health and health services*. UK: McGraw-Hill Education.
- Bradshaw, A. and Merriman, C. (2008) Nursing competence 10 years on: fit for practice and purpose yet? *Journal of clinical nursing*, 17(10), pp. 1263-69.
- Brannen, J. (2005) National Centre for Research Methods review papers, NCRM/005. Mixed methods research: A discussion paper. Available from: <http://eprints.ncrm.ac.uk/89/1/MethodsReviewPaperNCRM-005.pdf>. [Accessed 10th Feb 2017].

- Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), pp. 77-101.
- Bresler, L. (1995) Ethnography, phenomenology and action research in music education. *Quarterly Journal of Music Teaching and Learning*, 6(3), pp. 4-17.
- Brick, J.M. and Kalton, G. (1996) Handling missing data in survey research. *Statistical Methods in Medical Research*, 5(3), pp. 215-238.
- Broome, M.E. (1993) Integrative literature reviews for the development of concepts, in: Rodgers, B.L. and Knafl, K.A., (eds) *Concept Development in Nursing*. Philadelphia: W.B. Saunders Co, pp. 231-250.
- Brown L., Herd K. and Humphries, G. (2005) The role of the lecturer in practice placements: what do students think? *Nurse Education in Practice*, 5(2), pp. 84-90.
- Brown, S. (2005) Assessment for learning. *Learning and teaching in higher education*, (1), pp.81-89.
- Buck, A. J. Gross, M., Hakim, S. and Wainblatt, J. (1993) Using the Delphi process to analyse social policy implementation- a post hoc case of from vocational rehabilitation. *Policy Sciences*, 26(4), pp. 271-288
- Buckingham, S. (2000) Clinical competency: the right assessment tools? *Journal of Child Health Care*, 4(1), pp. 19-22.
- Burden, S., Topping, A.E. and O'Halloran, C. (2018) Mentor judgements and decision-making in the assessment of student nurse competence in practice: A mixed-methods study. *Journal of advanced nursing*, 74(5), pp. 1078-1089.
- Burgess, H. and Spangler, B. (2003) Consensus Building. A Beyond Intractability.org knowledge base essay. Available from: http://www.beyonدينtractability.org/essay/consensus_building [Accessed 5th December 2016]
- Burkhart-Kriesel, C., Trauman, K. and Rice, T. (2011) Benchmarking to Build. Sustainable Communities: "Expect, Like and Love to See". University of Nebraska. Available from: <http://extensionpublications.unl.edu/assets/pdf/ec860.pdf>. [Accessed 26th December 2016].
- Burnard, P., Gill, P., Stewart, K., Treasure, E. and Chadwick, B. (2008) Analysing and presenting qualitative data. *British dental journal*, 204(8), pp. 429-432.
- Butler, M.P., Cassidy, I., Quillinan, B., Fahy, A., Bradshaw, C., Tuohy, D., O'Connor, M., Mc Namara, M.C., Egan, G. and Tierney, C. (2011) Competency assessment methods—tool and processes: a survey of nurse preceptors in Ireland. *Nurse Education in Practice*, 11(5), pp. 298-303.
- Calman, L., Watson, R., Norman, I., Redfern, S. and Murrells, T. (2002) Assessing practice of student nurses: methods, preparation of assessors and students views. *Journal of Advanced Nursing*, 38(5), pp. 516-523.
- Cangelosi, P.R., Crocker, S. and Sorrell, J. M. (2009) Expert to Novice: Clinicians learning new roles as clinical educators. *Nursing Education Research*, 30(6), pp. 367-371.
- Cantrill, J.A., Sibbald, B. and Buetwo, S. (1996) The Delphi and nominal group techniques in health services research. *International Journal of Pharmacy Practice*, 4(2), pp. 67-74.
- Carney, O., McIntosh, J. and Worth, A. (1996) The use of nominal group technique in research with community nurses. *Journal of Advanced Nursing*, 23(5), pp. 1024-1029.

- Carpenter, C., Glasper, A., Jowett, R. and Nicholls, P. (2012) Celebrating 30 years of integrated undergraduate nursing at The University of Southampton. *Working Papers in Health Sciences*, 1(2), pp. 1-6.
- Cassidy, I., Butler, M.P., Quillinan, B., Egan, G., Mc Namara, M.C., Tuohy, D., Bradshaw, C., Fahy, A., O'Connor, M. and Tierney, C. (2012) Preceptors' views of assessing nursing students using a competency based approach. *Nurse Education in Practice*, 12(6), pp. 346-351.
- Cassidy, S. (2009) Subjectivity and the valid assessment of pre-registration student nurse clinical learning outcomes: implications for mentors. *Nurse Education Today*, 29(1), pp. 33-39.
- Cassidy, S., Coffey, M. and Murphy, F. (2017) 'Seeking authorization': a grounded theory exploration of mentors' experiences of assessing nursing students on the borderline of achievement of competence in clinical practice. *Journal of advanced nursing*, 73(9), pp. 2167-2178.
- Cavanagh, S. (1997) Content analysis: concepts, methods and applications. *Nurse researcher*, 4(3), pp. 5-13.
- Chandan, M. and Watts, C. (2012) Mentoring and pre-registration nurse education. Available from: <http://www.williscommission.org.uk> [Accessed 16th October 2016].
- Chong, E.J.M., Lim, J.S.W., Liu, Y., Lau, Y.Y.L. and Wu, V.X. (2016) Improvement of learning domains of nursing students with the use of authentic assessment pedagogy in clinical practice. *Nurse education in practice*, 20(September), pp.125-130.
- Clarke, A. and Dawson, R. (1999) *Evaluation research: An introduction to principles, methods and practice*. London: Sage.
- Clarke, C.L., Gibb, C.E. and Ramprogus, V. (2003) Clinical learning environments: an evaluation of an innovative role to support preregistration nursing placements. *Learning in health and social care*, 2(2), pp.105-115.
- Cleland, J.A., Knight, L.V., Rees, C.E., Tracey, S. and Bond, C.M. (2008) Is it me or is it them? *Medical Education*, 42(8), pp. 800-809.
- Clifford, C. (1994) Assessment of clinical practice and the role of the nurse teacher. *Nurse Education Today*, 14(4), pp. 272-279.
- Clutterbuck, D. (2004) Making the most of informal mentoring: A positive climate is key. *Development and Learning in Organizations: An International Journal*, 18(4), pp.16-17.
- Cockett, A. and Jackson, C. (2018) The use of assessment rubrics to enhance feedback in higher education: An integrative literature review. *Nurse education today*, 69(June), pp. 8-13.
- Cohen, G.S., Blumberg, P., Ryan, N.C and Sullivan, P.L. (1993) Do final grades reflect written qualitative evaluations of student performance? *Teaching and Learning in Medicine: An International Journal*, 5(1), pp. 10-15.
- Cooney, F., Stebbings S.N., Roxburgh, M., Mayo, J., Keen, N., Evans, E. and Meehan, T.C (1995) Integrating Nursing Research and Practice, Part 11 - a Delphi study of nursing practice priorities for research-based solutions. *Nursing Practice in New Zealand*, 10(1), pp. 22-27.
- Cooper, H.M. (1989) *Integrating research: A guide for literature reviews*, 2nd edn. CA: Sage Publications.
- Cooper, H.M. (1982) Scientific guidelines for conducting integrative research reviews. *Review of educational research*, 52(2), pp. 291-302.

- Cooper, H. M. (1998) *Synthesizing Research: A Guide for Literature Reviews*, 3rd edn. CA: Sage Publications.
- Coughlin, M., Cronin, P. and Ryan F. (2009) Survey Research: Process and Limitations. *International Journal of Therapy and Rehabilitation*, 16(1), pp. 9-15.
- Cowan, D.T., Norman, I. and Coopamah, V. P. (2005) Competence in nursing practice: A controversial concept – A focused review of literature. *Nurse education today*, 25(5), pp. 355-362.
- Cowan, D.T., Wilson-Barnett, J. and Norman, I.J. (2007) A European survey of general nurses' self assessment of competence. *Nurse Education Today*, 27(5), pp. 452-458.
- Crawford, C.L. and Johnson, J.A. (2012) To aspire or not: an integrative review of the evidence. *Nursing in Critical Care*, 7(5), pp. 9-15.
- Creswell J. W. (1997). *Qualitative inquiry and research design: Choosing among five traditions*. CA: Sage Publications.
- Creswell, J.W. and Clark, V.L.P (2017) *Designing and Conducting Mixed Methods Research*. 3rd Edn. CA: Sage Publications.
- Creswell, J.W., Fetters, M.D. and Ivankova, N.V. (2004) Designing a mixed methods study in primary care. *The Annals of Family Medicine*, 2(1), pp. 7-12.
- Crisp, B., Anderson, M., Orme, J. and Green Lister, P. (2003) *Learning and teaching in social work education: assessment*. Volume 1 of knowledge review (Vol. 1). Policy Press.
- Crossetti, M.G.O. (2012) Integrative review of nursing research: scientific rigour required. Available from: <http://www.scielo.br/pdf/rgenf/v33n2/03.pdf> [Accessed 24th June 2015].
- Crotty, M. (1993) The influence of educational theory on the development of nurse training to education in the United Kingdom. *Journal of advanced nursing*, 18(10), pp. 1645-1650.
- Daley, B.J. (1999) Novice to expert: An exploration of how professionals learn. *Adult education quarterly*, 49(4), pp. 133-147.
- Dalkey, N.C. (1969) The Delphi Method: an experimental study of group opinion. Document No RM-5888-PR. The RAND Corporation, Santa Monica, CA. Available from: https://www.rand.org/content/dam/rand/pubs/research_memoranda/2005/RM5888.pdf accessed 27/01/2016. [Accessed 1st January 2016].
- Darwin, C. (1959) *The Origin of Species by Means of Natural Selection*. London: J. J. Murray.
- Dawson, K.P. (2006) Jigsaws and Jugglers: Disposition, discourse and decision making in the assessment of student nurse practice. PhD Thesis, University of Stirling.
- Deegan, M.J. (1988) *Jane Addams and the Men of the Chicago School, 1892-1918*. New Brunswick, NJ: Transaction Books.
- Denzin, N. K. (1970) *The Research Act: A Theoretical Introduction to Sociological Methods*. New York: McGraw Hill.
- Department of Health (1999) Making a difference: strengthening the nursing, midwifery and health visiting contribution to health and health care. London: The stationery office.
- Dewey, J. (1916) *Democracy and Education*. New York: Macmillan.
- Dewey (1966) *Democracy and education: an introduction to the philosophy of education*. New York: Free Press.
- Dewey, J. (1909) *Moral Principles of Education*. Boston: Houghton Mifflin.
- Dewey, J. (1948) *Reconstruction in Philosophy*, Enlarged edn. Boston: The Beacon Press

- Dochy, F. J. and Segers, M. S. (1990) *Management information and performance indicators in higher education: An international issue*. Assen Maastricht: Van Gorcum.
- Dolan, G. (2003) Assessing student nurse clinical competency: will we ever get it right? *Journal of Clinical Nursing*, 12(1), pp. 132-141.
- Donaldson, J.H. and Gray, M. (2012) Systematic review of grading practice: is there evidence of grade inflation? *Nurse Education in Practice*, 12(2), pp. 101-114.
- Donohoe, H.M. and Needham, R.D. (2009) Moving best practice forward: Delphi characteristics, advantages, potential problems, and solutions. *International Journal of Tourism Research*, 11(5), pp. 415-437.
- Downe-Wamboldt, B. (1992) Content analysis: method, applications, and issues. *Health care for women international*, 13(3), pp. 313-321.
- Downing, S. M. (2005) Threats to the validity of clinical teaching assessments: What about rater error? *Medical Education*, 39(4), pp. 353-355.
- Dreyfus, S.E. and Dreyfus, H.L. (1980) *A five-stage model of the mental activities involved in directed skill acquisition* (No. ORC-80-2). California University Berkeley Operations Research Centre.
- Dudek, N. L., Marks, M. B. and Regehr, G. (2005) Failure to fail: the perspectives of clinical supervisors. *Academic Medicine: Journal of the Association of American Medical Colleges*, 80(10), pp. S84-S87.
- Duffy, K. (2013) Deciding to fail. *Journal of Practice Teaching and Learning*, 11 (3), pp. 36-58.
- Duffy, K. (2003) Failing students. A qualitative study of factors that influence the decisions regarding assessment of students' competence in practice London; NMC, London. Available from: http://www.nmcuk.org/Documents/Archived%20Publications/1Research%20papers/Kathleen_Duffy_Failing_Students2003.pdf [Accessed 11th October 2016].
- Duffy, K. and Watson, H.E. (2001) An interpretive study of the nurse teacher's role in practice placement areas. *Nurse Education Today*, 21(7), pp. 551-558.
- Earnshaw, G.J. (1995) Mentorship: the students' views. *Nurse Education Today*, 15(4), pp. 274-279.
- Ecken, P., Gnatzy, T. and Heiko, A. (2011) Desirability bias in foresight: consequences for decision quality based on Delphi results. *Technological Forecasting and Social Change*, 78(9), pp. 1654-1670.
- Elkan, R., Hillman, R. and Robenson, J. (1993) The implementation of Project 2000 in a District Health Authority; The effect on the Nursing service. A second interim report. University on Nottingham, Nottingham.
- Elo, S. and Kyngäs, H. (2008) The qualitative content analysis process. *Journal of advanced nursing*, 62(1), pp.107-115.
- English National Board (1989) Project 2000 - A new preparation for practice Guidelines and criteria for course development English National Board for Nursing, Midwifery and Health Visiting, London.
- English National Board for Nursing Midwifery and Health Visiting and Department of Health (2001a) Placements in Focus. ENB, London.
- English National Board for Nursing Midwifery and Health Visiting and Department of Health (2001b) Preparation of Mentors and Teachers. ENB, London.
- Eno, S. and Kerr, J. (2013) 'That was awful! I'm not ready yet, am I?' Is there such a thing as a Good Fail? *The Journal of Practice Teaching and Learning*, 11(3), pp. 135-148.

- Enrico, N.B. and Chapman, Y. (2011) The Lived Experience of Mentoring in Malaysia. *Nurse Media Journal of Nursing*, 1(1), pp. 87-104.
- Eraut, M. (1994) *Developing professional knowledge and competence*. London: Falmer Press.
- Erffmeyer, R.C., Erffmeyer, E.S. and Lane, I.M. (1986) The Delphi technique: An empirical evaluation of the optimal number of rounds. *Group and Organization Management*, 11(1-2), pp. 120-128.
- Evans, C. (1997) The use of consensus methods and expert panels in pharmacoeconomic studies, practical applications and methodological shortcomings. *Pharmacoeconomics*, 12(2), pp. 121-129.
- Fahy, A., Tuohy, D., McNamara, M.C., Butler, M.P., Cassidy, I. and Bradshaw, C. (2011) Evaluating clinical competence assessment. *Nursing Standard*, 25(50), pp. 42-48.
- Farrimond, H. (2013) *Doing Ethical Research*. Basingstoke: Macmillian.
- Finch, J., Poletti, A. (2013) "It's been hell" Italian and British Practice Assessors' Narratives of Working with Struggling or Failing Social Work Students in Practice Learning Settings. *European Journal of Social Work*, 17(1), pp. 135-150.
- Finch, J., Taylor, I. (2013) Failure to Fail? Practice Educators' Emotional Experiences of Assessing Failing Social Work Students. *Social Work Education*, 32(2), pp. 244-258.
- Finfgeld-Connett, D. (2014) Use of content analysis to conduct knowledge building and theory-generating qualitative systematic reviews. *Qualitative Research*, 14(3), pp. 341-352.
- Fish, J. (2018) *Pan London Practice Assessment Document*. In *Stakeholder Event presentation*, Camden- London, 26 February, pp. 10.
- Fitzgerald, M., Gibson, F. and Gunn, K. (2010) Contemporary issues relating to assessment of pre-registration nursing students in practice. *Nurse education in practice*, 10(3), pp. 158-163.
- Fotheringham, D. (2011) The role of expert judgement and feedback in sustainable assessment: A discussion paper. *Nurse education today*, 31(8), pp. e47-e50.
- Fowler Jr, F.J. (2013) *Survey research methods*. London: Sage.
- Freire, P. (1970) *Pedagogy of the oppressed*. New York: Herder and Herder.
- Frentsos, J.M. (2013) Rubrics role in measuring nursing staff competencies. *Journal for nurses in professional development*, 29(1), pp.19-23.
- Fulbrook, P., Rolfe, G., Albarran, J. and Boxall, F. (2000) Fit for practice: Project 2000 student nurses' views on how well the curriculum prepares them for clinical practice. *Nurse Education Today*, 20(5), pp.350-357.
- Fulton, J., Böhler, A., Hansen, G.S., Kauffeldt, A., Welander, E., Santos, M.R., Thorarinsdottir, K. and Ziarko, E. (2007) Mentorship: An international perspective. *Nurse Education in practice*, 7(6), pp. 399-406.
- Gainsbury, S. (2010) Mentors admit to passing bad students. *Nursing Times*, 106(16), pp. 1-3.
- Gallagher, P., Smith T. and Ousey, K. (2012) Problems with competence assessment as it applies to student nurses. *Nurse education in practice*, 12(6), pp. 301-303.
- Garside, J.R. and Nhemachena, J.Z. (2013) A concept analysis of competence and its transition in nursing. *Nurse Education Today*, 33(5), pp. 541-545.
- Gerrish, K., McManus, M. and Ashworth, P. (1997) *Levels of Achievement: A Review of the Assessment of Practice*. *Researching Professional Education. Research Reports Series*

- Number 5. English National Board for Nursing, Midwifery and Health Visiting, Publications.
- Gibbs, G. (1988) *Learning by doing: A guide to teaching and learning methods*. London: Further Education Unit.
- Gibbs, G. and Simpson, C. (2005) Conditions under which assessment supports students' learning. *Learning and teaching in higher education*, (1), pp. 3-31.
- Girof, E.A. (2000) Assessment of graduates and diplomates in practice in the UK—are we measuring the same level of competence? *Journal of Clinical Nursing*, 9(3), pp. 330-337.
- Goodman, C. M. (1987). The Delphi technique: a critique. *Journal of advanced nursing*, 12(6), pp. 729-734.
- Gopee, N. (2015) *Mentoring and supervision in health care*. 3rd edn. London: Sage.
- Govaerts, M.J.B. and Van der Vleuten, C.P.M. (2013) Validity of work-based assessment: exploring our horizons. *Medical Education*, 47(12), pp. 1164-1174.
- Govaerts, M.J.B., Van de Wiel, M.W.J., Schuwirth, L.W.T., Van der Vleuten, C.P.M. and Muijtjens, A.M.M. (2013) Workplace-based assessment: raters' performance theories and constructs. *Advances in Health Sciences Education*, 18(3), pp. 375-396.
- Gray, D.P. (1993) Assessment at last. *The British Journal of General Practice*, 43(375), pp. 402-403.
- Gray, M.A. and Donaldson, J. (2009) Exploring issues in the use of grading in practice; literature review. *Edinburgh Napier University, Edinburgh*.
- Gray, J.R., Grove, S.K. and Sutherland, S. (2017) *Burns and Grove's The Practice of Nursing Research: Appraisal, Synthesis, and Generation of Evidence*. 8th edn. St. Louis, MO: Elsevier Health Sciences.
- Grealish, L.A. and Shaw, J.M. (2018) Introducing a design exigency to promote student learning through assessment: A case study. *Nurse education today*, 61(February), pp. 225-230.
- Greene, J. (2008) "Is Mixed Methods Social Inquiry a Distinctive Methodology." *Journal of Mixed Methods Research*, 2(7), pp. 7-22.
- Greene, J. and McClintock, C. (1985) Triangulation in evaluation: Design and analysis issues. *Evaluation Review*, 9(5), pp. 523-545.
- Guba, E.G. and Lincoln, Y.S. (2005) Paradigmatic controversies, contradictions, and emerging confluences. *The Sage handbook of qualitative research*, 3rd ed. CA: Sage.
- Gulikers, J.T., Bastiaens, T.J. and Kirschner, P.A. (2004) A five-dimensional framework for authentic assessment. *Educational technology research and development*, 52(3), pp. 67-86.
- Habermas, J. (1984) *The theory of communicative action, volume 1: reason and the rationalisation of society*. Boston, MA: Beacon.
- Haines, C., Dennick, R. and da Silva, J.A.P. (2013) Developing a professional approach to work-based assessments in rheumatology. *Best Practice and Research Clinical Rheumatology*, 27(2), pp. 123-136.
- Haitana, J. and Bland, M. (2011) Building relationships: the key to preceptoring nursing students. *Nursing Praxis in New Zealand*, 27(1), pp. 4-12.
- Halpin, Y., Terry, L.M. and Curzio, J. (2017) A longitudinal, mixed methods investigation of newly qualified nurses' workplace stressors and stress experiences during transition. *Journal of advanced nursing*, 73(11), pp. 2577-2586.

- Hardy, E.C., Koharchik, L.S. and Dixon, H. (2015) The professional nurse–student nurse academic partnership. *Teaching and Learning in Nursing*, 10(2), pp. 71-75.
- Harvey, L. and Green, D. (1993) Defining quality. *Assessment and evaluation in higher education*, 18(1), pp. 9-34.
- Hasson, F., Keeney, S. and McKenna, H. (2000) Research guidelines for the Delphi survey technique. *Journal of advanced nursing*, 32(4), pp. 1008-1015.
- Hatala, R. and Norman, G.R. (1999) In-training evaluation during an internal medicine clerkship. *Academic Medicine*, 74(10), pp. S118–S120.
- Hawkins, K. (2008) Participatory action research, sacred existential epistemology, the eighth moment of qualitative research and beyond.... In *Proceedings of the AARE Conference 2007: Research Impacts-Proving or Improving?* Australian Association for Research in Education.
- Hayes, K. (2010) Work-place assessment. *Obstetrics, Gynaecology and Reproductive Medicine*, 21(2), pp. 52-54.
- Hays, R. and Wellard, R. (1998) In-training assessment in postgraduate training for general practice. *Medical education*, 32(5), pp. 507-513.
- Health Education England (2018) Reducing Pre-Registration Attrition and Improving Retention (RePAIR) Project. Available from: <https://www.hee.nhs.uk/our-work/reducing-pre-registration-attrition-improving-retention> [Accessed 21 October 2018].
- Heaslip, V. and Scammell, J.M. (2012) Failing underperforming students: the role of grading in practice assessment. *Nurse Education in practice*, 12(2), pp. 95-100.
- Helminen, K., Coco, K., Johnson, M., Turunen, H. and Tossavainen, K. (2016) Summative assessment of clinical practice of student nurses: A review of the literature. *International journal of nursing studies*, 53(January), pp. 308-319.
- Hesse-Biber, S. (2010) Qualitative approaches to mixed methods practice. *Qualitative Inquiry*, 16(6), pp. 455-468.
- Hicks, C. (2009) *Research Methods for clinical therapists applied project design and analysis*. 5th edn. London: Churchill Livingstone.
- Hill, K.Q. and Fowles, J. (1975) The methodological worth of the Delphi forecasting technique. *Technological Forecasting and Social Change*, 7(2), pp. 179-192.
- Howe, E. (2003) It's pretty difficult to fail. *Assessment in higher education*, 28(4), pp. 371-382.
- Hoyer, R. W. and Hoyer, B. B. Y. (2001) What is quality? *Quality Progress*, 34(7), pp. 53–62.
- Hsieh, H.F. and Shannon, S.E. (2005) Three approaches to qualitative content analysis, *Qualitative health research*, 15(9), pp. 1277-1288.
- Hsu, C.C. and Sandford, B.A. (2007) The Delphi technique: making sense of consensus. *Practical assessment, research and evaluation*, 12(10), pp.1-8.
- Hunt, L.A. (2014) Failing Securely: Enabling mentors to fail underperforming student nurses in practical assessments. Doctoral dissertation, Birmingham City University.
- Hunt, L.A., McGee, P., Gutteridge, R. and Hughes, M. (2012) Assessment of student nurses in practice: A comparison of theoretical and practical assessment results in England. *Nurse Education Today*, 32(4), pp. 351-355.
- Hunt, L.A., McGee, P., Gutteridge, R. and Hughes, M. (2016) Failing securely: The processes and support which underpin English nurse mentors' assessment decisions regarding under-performing students. *Nurse education today*, 39(April), pp. 79-86.

- Ihantola, E.M. and Kihn, L.A. (2011) Threats to validity and reliability in mixed methods accounting research. *Qualitative Research in Accounting and Management*, 8(1), pp. 39-58.
- Ilott, I. and Murphy, R. (1997) Feelings and failing in professional training: the assessor's dilemma. *Assessment and Evaluation in Higher Education*, 22(3), pp. 307-316.
- Jairath, N., and Weinstein, J. (1994) The Delphi Methodology: a useful administrative approach. *Canadian Journal of Nursing Administration*, 7(3), pp. 29-42.
- James, M. (2006) Assessment, teaching and theories of learning, in: Gardner, J. (Ed), *Assessment and learning*. London: Sage, pp. 47-60.
- James, W. (1907) Pragmatism, a new name for some old ways of thinking. *Pragmatism and Other Writings*. New York: Penguin.
- James, W. (1890) The consciousness of self. *The principles of psychology*. Available from: <http://cosmology.com/Consciousness123.html> [Accessed 4th January 2017].
- Jervis, A. and Tilki, M. (2011) Why are nurse mentors failing to student nurses who do not meet clinical performance standards? *British Journal of Nursing*, 20(9), pp. 582-587.
- Jinks, A. and Haroon-Iqbal, H. (2002) Peer review of clinical assessment: Results of a clinical mentor survey. *Singapore Nursing Journal*, 29(2), pp. 8-12.
- Johnson, R.B. and Onwuegbuzie, A.J. (2004) Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), pp. 14-26.
- Johnson, R.B., Onwuegbuzie, A.J. and Turner, L.A. (2007) Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), pp. 112-133.
- Jokelainen, M., Tossavainen, K., Jamookeeah, D. and Turnen, H. (2013) Finish and British mentors' conceptions of facilitating nursing students' placement learning and professional development. *Nurse Education in Practice*, 13(1), pp. 61-67.
- Jones, R.A. (2005) Prioritising to safeguard the integrity of nurse education. *Nursing Times*, 101(34), pp. 26-27.
- Jones, J. and Hunter, D. (1995) Consensus methods for medical and health services research. *British Medical Journal*, 311(7001), pp. 376-380.
- Jonsson, A. and Svingby, G. (2007) The use of scoring rubrics: reliability, validity and educational consequences. *Education research review*, 2(2), pp. 130-144.
- Jowett, S., Walton, I. and Payne, S. (1994) *Challenges and Changes in Nurse Education: A Study of the Implementation of Project 2000*. UK: National Foundation for Educational Research in England and Wales.
- Jowett, S., Walton, I. and Payne, S. (1992) *Implementing Project 2000: an interim report*. UK: National Foundation for Educational Research in England and Wales.
- Juran, J. and Godfrey, A.B. (1999) *Juran's Quality handbook*. New York: McGraw-Hill.
- Kadlec, A. (2007) *Dewey's Critical Pragmatism*. Plymouth: Lexington Books.
- Kadlec, A. (2006). Reconstructing Dewey: the philosophy of critical pragmatism. *Polity*, 34(4), pp. 519-542.
- Keeney, S., Hasson, F. and McKenna, H. (2006) Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of advanced nursing*, 53(2), pp. 205-212.
- Keeney, S. Hasson, F. and McKenna, H. (2011) *The Delphi Technique in Nursing and Health Research*. Chichester: Wiley-Blackwell.
- Kendrick, K.D. and Simpson, A. (1992) Project 2000: the nurse's reformation, in: Henry, C., Kendrick, K. and Soothill, K. (eds.) *Themes and perspectives in nursing*. London: Chapman and Hall, pp. 91-100.

- Khandker, S.R., Koolwal, G.B. and Samad, H.A. (2009) *Handbook on impact evaluation: quantitative methods and practices*. Washington DC: World Bank Publications.
- King, L. and Macleod Clark, J. (2002) Intuition and the development of expertise in surgical ward intensive care nurses, *Journal of Advanced Nursing*, 37(4), pp. 322-329.
- Koh, K. (2017) *Authentic Assessment*. Oxford Research Encyclopaedia of Education. Available from: <http://education.oxfordre.com/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-22>. [Accessed 15th October 2018].
- Krippendorff, K. (2012) *Content analysis: an introduction to its methodology*. CA: Sage.
- Krueger, R.A. and Casey, M.A. (2014) *Focus groups: A practical guide for applied research*. London: Sage.
- Landeta, J. (2006) Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, 73(5), pp. 467-82.
- Lang, T.A. and Secic, M. (2006) *How to report statistics in medicine: annotated guidelines for authors, editors, and reviewers*. American Collage of Physicians Press.
- Lankshear, A. (1990) Failure to fail: the teacher's dilemma, *Nursing Standard*, 4(20), pp. 35-37.
- Larocque, S. and Luhanga, F.L. (2013) Exploring the issue of failure to fail in a nursing program. *International Journal of Nursing Education Scholarship*, 10(1), pp. 1-8.
- Lathlean, J. (1989) *Policy making in nurse education*. St George's School of Nursing.
- Lave, J. and Wenger, E. (1991) *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge university press.
- Lawson, L. (2010) *Supporting mentors and clinical educators: a collaborative project into the development of knowledge and skills to enhance clinical education practice*. Unpublished report. University of Hertfordshire.
- Le Deist, F.D. and Winterton, J. (2005) What is competence? *Human resource development international*, 8(1), pp. 27-46.
- Lemmer, B. (1998) Successive surveys of an expert panel: research in decision-making with health visitors. *Journal of advanced nursing*, 27(3), pp. 538-545.
- Leonard, A., Jowett, S. (1990) Charting the Course: A Study of the 6 ENB Pilot Schemes in Pre-registration Nurse Education, in: *The National Evaluation of Demonstration Schemes in Pre-registration Nurse Education*. London: National Foundation for Educational Research in England and Wales.
- Levett-Jones, T., Gersbach, J., Arthur, C. and Roche, J. (2011) Implementing a clinical competency assessment model that promotes critical reflection and ensures nursing graduates' readiness for professional practice. *Nurse Education in Practice*, 11(1), pp. 64-69.
- Levett-Jones, T., Lathlean, J., Higgins, I. and McMillan, M. (2009) Staff–student relationships and their impact on nursing students' belongingness and learning. *Journal of advanced nursing*, 65(2), pp. 316-324.
- Licari, F. and Chambers, D.W. (2008) Some Paradoxes in Competency Based Dental Education. *Journal of Dental Education*, 72(1) pp. 8-18.
- Lincoln, Y.S. And Guba, E.G. (1985) *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications.
- Lincoln, Y.S. and Guba, E.G. (1986) Research, evaluation, and policy analysis: Heuristics for disciplined inquiry. *Review of Policy Research*, 5(3), pp. 546-565.

- Lindop, E. (1989) Individual stress and its relationship to termination of nurse training. *Nurse Education Today*, 9(3), pp. 172-179.
- Linell, P. (2009) *Rethinking language, mind, and world dialogically*. Charlotte, North Carolina: Information Age Publishing.
- Linstone H.A. (1978) The Delphi technique. In Handbook of Futures Research, in: Fowles, R.B. (ed.) *Handbook of futures research*. Connecticut: Greenwood Press, pp. 273–300.
- Linstone, H.A. and Turoff, M. (2011) Delphi: A brief look backward and forward. *Technological Forecasting and Social Change*, 78(9), pp. 1712-1719.
- Linstone, H. A and Turoff, M. (1975) *The Delphi Method: Techniques and Applications*. Massachusetts: Addison – Wesley.
- Llaurado-Serra, M., Rodríguez, E., Gallart, A., Fuster, P., Monforte-Royo, C. and De Juan, M.Á. (2018) Assessing the competences associated with a nursing Bachelor thesis by means of rubrics. *Nurse education today*, 66(July), pp. 103-109.
- Lock, L. R. (2011) Selecting examinable nursing core competencies: a Delphi project. *International nursing review*, 58(3), pp. 347-353.
- Longley, M., Shaw, C. and Dolan, G. (2007) Alternative Scenarios for Healthcare, Nursing and Nurse Education in the UK in 2015. Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.453.8839> [Accessed 17th July 2016].
- Loughlin, K.G. and Moore, L.F. (1979) Using Delphi to achieve congruent objectives and activities in a paediatrics department. *Journal of medical education*, 54(2), pp.101-106.
- Luhanga, F., Yonge, O.j. and Myrick, F. (2008a) Precepting an unsafe student: the role of the faculty. *Nurse Education Today*, 28(2), pp. 227-231.
- Luhanga, F., Yonge, O.J. and Myrick, F. (2008b) Strategies for precepting the unsafe student. *Journal for nurses in staff development*, 24(5), pp. 214-219.
- Luhanga, F., Yonge, O.J. and Myrick, F. (2008c) " Failure to Assign Failing Grades": Issues with Grading the Unsafe Student. *International Journal of Nursing Education Scholarship*, 5(1), pp. 1-14.
- MacDougall, M., Riley, S.C., Cameron, H.S. and McKinstry, B. (2008) Halos and Horns in the Assessment of Undergraduate Medical Students: A Consistency-Based Approach. *Journal of applied quantitative methods*, 3(2), pp.116-128.
- Maguire P. (1987) *Doing Participatory Research*. Massachusetts: University of Massachusetts, Amherst.
- Mason, J. (2006) 'Six strategies for mixing methods and linking data in social science research', Real Life Methods working paper. University of Manchester.
- May, V. (2010) Realities Toolkit 11: 'What to do with contradictory data?' Available from: <http://eprints.ncrm.ac.uk/1322/1/12-toolkit-contradictory-data.pdf>. [Accessed 26th Feb 2017].
- McCarthy, B. and Murphy, S. (2008) Assessing undergraduate nursing students in clinical practice: do preceptors use assessment strategies? *Nurse Education Today*, 28(3), pp. 301-313.
- McClellan, D. (1976) Testing for Competence Rather than 'Intelligence.' Available from: <https://www.therapiebreve.be/documents/mcclelland-1973.pdf> [Accessed 12th January 2018].

- McDermott, M.Z.R., Daniels, M., Nylén, A., Pears, A., Isomöttönen, V. and Caspersen, M. (2017) The authenticity of 'authentic' assessment some faculty perceptions, in: *Frontiers in Education Conference*, pp. 1-9. Available from: <https://jyx.jyu.fi/bitstream/handle/123456789/56812/1/authenticity.pdf>. [Accessed 12th May 2018].
- McDowell, L., Wakelin, D., Montgomery, C. and King, S. (2011) Does assessment for learning make a difference? The development of a questionnaire to explore the student response. *Assessment and Evaluation in Higher Education*, 36(7), pp. 749-765.
- McKenna, H. P. (1994) The Delphi technique: a worthwhile approach for Nursing? *Journal of Advanced Nursing*, 19(6), pp. 1221-1225.
- McMullan, M., Endacott, R., Gray, M.A., Jasper, M., Miller, C.M., Scholes, J. and Webb, C. (2003) Portfolios and assessment of competence: a review of the literature. *Journal of advanced nursing*, 41(3), pp. 283-294.
- Mead, D., Hopkins, A. and Wilson, C. (2011) Views of nurse mentors about their role. *Nursing Management*, 18(6), pp. 18-23.
- Menand, L. (1997) *Pragmatism: A reader*. New York: Vintage Books.
- Messick, S. (1995) Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, 50(9), pp.741-749.
- Miller, C. (2010) Improving and enhancing performance in the affective domain of nursing students: Insights from the literature for clinical educators. *Contemporary Nurse*, 35(1), pp. 2-17.
- Mitchell, M.P. (1998) Nursing education planning: a Delphi study. *Journal of Nursing Education*, 37(7), pp. 305-307.
- Moher, D., Liberati, A., Tetzlaff, J. and Altman, D.G. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), pp. 264-269.
- Montgomery, K. (2002) Authentic Tasks and Rubrics: going beyond traditional assessments in College Teaching. *College teaching*, 50(1), pp. 34-40.
- Moore, D. (2005) Assuring Fitness for Practice, A Policy Review Commissioned by the Nursing Midwifery Council Nursing Task and Finish Group. London: Nursing Midwifery Council.
- Morse, J.M., Barrett, M., Mayan, M., Olson, K. and Spiers, J. (2002) Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, 1(2), pp. 13-22.
- Moskal, B.M. and Leydens, J.A. (2000) Scoring rubric development: Validity and reliability. *Practical assessment, research and evaluation*, 7(10), pp. 71-81.
- Moule, P., Aveyard, H. and Goodman, M. (2017) *Nursing research: An introduction*. 3rd edn, London: Sage.
- Murphy, M.K., Black, N., Lamping D.L., McKee, C.M. and Sanderson, C.F.B. (1998) Consensus development methods and their use in clinical guideline development. *Health Technology Assessment*, 2(3), pp. 1-88
- Murray, C. and Staniland, K. (2010) Mentorship: past, present and future, in: *The Nurse Mentor and Reviewer Update Handbook*. Maidenhead: Open University Press, pp.3-16.

- Murry Jr, J.W. and Hammons, J.O. (1995) Delphi: A versatile methodology for conducting qualitative research. *The Review of Higher Education*, 18(4), pp. 423-436.
- Neary, M. (1996) Continuous assessment of clinical competence: an investigation. Unpublished PhD thesis, School of Education, University of Wales, Cardiff.
- Neary, M. (2000a) Responsive assessment of clinical competence: Part 1. *Nursing Standard*, 15(9), pp. 34-36.
- Neary, M. (2000b) Responsive assessment of clinical competence: Part 2. *Nursing Standard*, 15(10), pp. 35-40.
- Neary, M. (2001) Responsive assessment: assessing student nurses' clinical competence. *Nurse Education Today*, 21(1), pp. 3-17.
- Nettleton, P. and Bray, L. (2008) Current mentorship schemes might be doing our students a disservice. *Nurse Education in Practice*, 8(3), pp. 205-212.
- NHS Employers (2014) NHS Qualified Nurse Supply and Demand Survey—Findings. Report produced for the Health Education England Nursing Supply Steering Group. NHS Employers, Leeds. Available from: <http://www.nhsemployers.org/-/media/Employers/Publications/NHS-qualified-nurse-supply-and-demand-survey-12-May.pdf> [Accessed 25th May 2018].
- NHS Modernisation Agency (2004) *10 High Impact Changes for Service Improvement and Delivery: A Guide for NHS Leaders*. NHS Modernisation Agency.
- Norman, I.J., Watson, R., Murrells, T., Calman, L. and Redfern, S. (2002) The validity and reliability of methods to assess the competence to practise of pre-registration nursing and midwifery students. *International Journal of Nursing Studies*, 39(2), pp. 133-145.
- Nursing and Midwifery Council (2004b) Consultation on a standard to support learning and assessment in practice. NMC: London.
- Nursing and Midwifery Council (2017) Consultation on education framework: standards for education and training, London, NMC. [www] available from: <https://www.nmc.org.uk/education/education-consultation/framework> [accessed 19th June 2017].
- Nursing and Midwifery Council (2005) Consultation on fitness for practice at the point of registration. NMC: London.
- Nursing and Midwifery Council (2018c) Nursing and Midwifery Council Annual Fitness to Practise Report 2017-2018. NMC: London.
- Nursing and Midwifery Council (2006) Standards for Support Learning and Assessment in Practice. NMC: London.
- Nursing and Midwifery Council (2004a) Standards of proficiency for pre-registration nursing education. NMC: London.
- Nursing and Midwifery Council (2008a) Standards to Support Learning and Assessment in Practice: NMC Standards for Mentors, Practice Teachers and Teachers. NMC: London.
- Nursing and Midwifery Council (2010) Standards for pre-registration nursing education. NMC: London.
- Nursing and Midwifery Council (2008b) The Code: Standards of conduct, performance and ethics for nurses and midwives. NMC: London.
- Nursing and Midwifery Council (2015) The Code: Professional standards of practice and behaviour for nurses and midwives. NMC: London.

- Nursing and Midwifery Council (2018d) The Code: Professional standards of practice and behaviour for nurses and midwives. NMC: London.
- Nursing and Midwifery Council (2018a) Standards for student supervision and assessment. NMC: London.
- Nursing and Midwifery Council (2018b) Standards for pre-registration nursing programmes. NMC: London.
- Nursing and Midwifery Council (2018e) What can 'support' mean? London, NMC. [www] available from: <https://www.nmc.org.uk/supporting-information-on-standards-for-student-supervision-and-assessment/practice-assessment/support-for-the-practice-assessor-role/what-can-support-mean/> [accessed 20th January 2019].
- Onwuegbuzie, A.J. and Collins, K.M. (2007) A typology of mixed methods sampling designs in social science research. *The qualitative report*, 12(2), pp. 281-316.
- Onwuegbuzie, A. J. (2003) Expanding the framework of internal and external validity in quantitative research. *Research in the Schools*, 10(1), pp. 71-90.
- Onwuegbuzie, A.J. and Johnson, R.B. (2006) The validity issue in mixed research. *Research in the Schools*, 13(1), pp. 48-63.
- Onwuegbuzie, A.J. and Teddlie, C. (2003) A framework for analyzing data in mixed methods research, in: A.Tashakkori, A. and Teddlie, C. (Eds.) *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage, pp. 351-383.
- Oppenheim, A.N. (1992) *Questionnaire design, interviewing and attitude measurement*. New York: Bloomsbury Publishing.
- Orr, J. M., Sackett, P. R., and DuBois, C. L. Z. (1991) Outlier detection and treatment in I/O Psychology: A survey of researcher beliefs and an empirical illustration. *Personnel Psychology*, 44(3), pp. 473-486.
- Osborne, J.W. and Overbay, A. (2004) The power of outliers (and why researchers should always check for them). *Practical assessment research and evaluation*, 9(6), pp. 1-12
- Ousey, K. (2011) The changing face of student nurse education and training programmes. *Wounds UK*, 7(1), pp.70-75.
- Owens C., Ley A. and Aitken P. (2008) Do different stakeholder groups shape mental health service priorities? A four arm Delphi Study. *Health Expectations*, 11(4), pp. 418-431.
- Oxford Centre for Evidence-Based Medicine (2009) Levels of Evidence. Available from: <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/> [Accessed 4th June 2016].
- Oxford English dictionary (2018) available from: <https://en.oxforddictionaries.com/definition/assess> [Accessed 15th December 2018]
- Pace, R., Pluye, P., Bartlett, G., Macaulay, A.C., Salsberg, J., Jagosh, J. and Seller, R. (2012) Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. *International journal of nursing studies*. 49(1), pp. 47-53.
- Paisley, A. M., Baldwin, P. J. and Paterson-Brown, S. (2005) Accuracy of medical staff assessment of trainees' operative performance, *Medical Teacher*, 27(7), pp. 634-638.
- Parratt, J.A., Fahy, K.M., Hutchinson, M., Lohmann, G., Hastie, C.R., Chaseling, M. and O'Brien, K. (2016) Expert validation of a teamwork assessment rubric: A modified Delphi study. *Nurse education today*, 36(January), pp. 77-85.
- Patton, M.Q. (2002) *Qualitative interviewing and research methods*, 2nd edn. London: Sage.

- Payne, S., Jowett, S. and Walton, I. (1991) *Nurse Teachers in Project 2000: The Experience of Planning and Initial Implementation*. Slough, Berkshire: National Foundation for Educational Research in England and Wales.
- Peirce, C.S. (1877) *The fixation of belief*. Available from: <http://bocc.ubi.pt/pag/peirce-charles-fixation-belief.html> [Accessed 2nd November 2015].
- Pellatt, G.C. (2006) The role of mentors in supporting pre-registration nursing students. *British Journal of Nursing*, 15(6), pp. 336–40.
- Penciner, R., Langan, T., Lee, R., McEwen, J., Woods, R. A. and Bandiera, G. (2011) Using a Delphi process to establish consensus on emergency medicine clerkship competencies. *Medical teacher*, 33(6), pp. e333-e339.
- Phillips, R.M., Davies, W.B. and Neary, M. (1996) The practitioner-teacher: a study in the introduction of mentors in the preregistration nurse education programme in Wales: part 1. *Journal of Advanced Nursing*, 23(5), pp. 1037-1044.
- Plano Clarke, A. and Dawson, R. (1999) *Evaluation research: An introduction to principles, methods and practice*. London: Sage.
- Pluye, P., Gagnon, M.P., Griffiths, F. and Johnson-Lafleur, J. (2009) A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *International journal of nursing studies*, 46 (4), pp. 529–546.
- Polit, D.F. and Beck, C.T. (2014) *Essentials of nursing research*. Philadelphia: Lippincott Williams and Wilkins.
- Powell, C. (2003) The Delphi Technique: Myths and Realities. *Journal of Advanced Nursing*, 41(4), pp. 376-382.
- Punch, K.F. and Oancea, A. (2014) *Introduction to research methods in education*. CA: Sage.
- Quality Assurance Agency (2012) *The UK Quality Code for Higher Education*. Available from: <http://www.qaa.ac.uk/assuring-standards-and-quality/the-quality-code> [Accessed 25th August 2016].
- Rattray, J. and Jones, M.C. (2007) Essential elements of questionnaire design and development. *Journal of clinical nursing*, 16(2), pp. 234-243.
- Rawles, J. (2013) Whose students are they anyway? *The Journal of Practice Teaching and Learning*, 11(3), pp. 59-78.
- Reeves, G. and Jauch, L.R. (1978) Curriculum development through Delphi. *Research in Higher Education*, 8(2), pp. 157-168.
- Rennert-Ariev, P. (2005) A theoretical model for the authentic assessment of teaching. *Practical Assessment Research and Evaluation*, 10(2), pp. 151-163.
- Robbins, J., Firth, A. and Evans, M. (2018) Improving work based assessment: addressing grade inflation numerically or pedagogically? *Practitioner Research in Higher Education*, 11(1), pp. 80-86.
- Robinson, S., Cornish, J., Driscoll, C., Knutton, S., Corben, V. and Stevenson, T. (2012) *Sustaining and managing the delivery of student nurse mentorship: roles, resources, standards and debates*. National Nursing Research Unit. Kings College London. Available from: <http://www.kcl.ac.uk/nursing/research/nnru/publications/Reports/Nurse-Mentorship-Report-Nov2012.pdf> [Accessed 12 October 2015].
- Rogers, P., Hawkins, A., McDonald, B., Macfarlan, A. and Milne, C. (2015) *Choosing appropriate designs and methods for impact evaluation*. Office of the Chief

- Economist, Department of Industry, Innovation and Science. Australian Government.
- Rohrbaugh, J. (1979) Improving the quality of group judgement: Social judgement analysis and the Delphi technique. *Organizational Behavior and Human Performance*, 24(1), pp. 73-92.
- Rowe, G. and Wright, G. (1999) The Delphi technique as a forecasting tool: Issues and analysis. *International Journal of Forecasting*, 15(4), pp. 353–375.
- Rowe, G. and Wright, G. (1996) The impact of task characteristics on the performance of structured group forecasting techniques. *International Journal of Forecasting*, 12(1), pp.73-89.
- Rowe, G. and Wright, G. (2001) Expert opinions in forecasting: the role of the Delphi technique, in: Armstrong, J.S. (Ed.) *Principles of Forecasting: A Handbook of Researchers and Practitioners* (Vol. 30) USA: Springer Science and Business Media, pp. 125-144.
- Royal College of Nursing (1985) *The education of Nurse: A new Dispensation. The Judge report*. London: RCN.
- Russell, C.L. (2005) An overview of the integrative research review. *Progress in transplantation*, 15(1), pp. 8-13.
- Sabey, A. and Harris, M. (2012) 'It's the conversation they'll learn from': improving assessments for GP Specialist Trainees in hospital posts, Education for Primary Care: An Official Publication of the Association of Course Organisers, National Association of GP Tutors. *World Organisation of Family Doctors*, 23(4), pp. 263-269.
- Sackman, H. (1975) *Delphi Critique*. Boston Massachusetts: Lexington Books.
- Sadler, D.R. (1998) Formative assessment: Revisiting the territory. *Assessment in education: principles, policy and practice*, 5(1), pp. 77-84.
- Scheff, T.J. (2006) *Goffman unbound. A new paradigm for social science*. London: Routledge.
- Scholes, J., Freeman, M., Gray M., Wallis, M., Robinson, D., Matthews-Smith, G. and Miller, C. (2004) *Evaluation of Nurse Education Partnership*, Department of Health, London.
- Schön, D.A. (1983) *The reflective practitioner: how professionals think in action*. USA: Basic Books Inc.
- Schuwirth, L.W. and Van der Vleuten, C.P. (2011) Programmatic assessment: from assessment of learning to assessment for learning. *Medical teacher*, 33(6), pp. 478-485.
- Sharp, M., Danby, H. (2000) *The Management of Failing Diploma Social Work Students*. London: Ashgate.
- Shields, P.M. (1998) Pragmatism as a philosophy of science: A tool for public administration, in: White, J. *Research in Public Administration*. New York: JAI Press Inc., pp. 195-225.
- Shipman, D., Roa, M., Hooten, J. and Wang, Z.J. (2012) Using the analytic rubric as an evaluation tool in nursing education: the positive and the negative. *Nurse education today*, 32(3), pp. 246-249.
- Shuttleworth, M. (2009) *Repeated Measure Design*. Available from: <https://explorable.com/repeated-measures-design> [Accessed 7th May 2017].
- Skulmoski, G.J., Hartman, F.T. and Krahn, J. (2007) The Delphi method for graduate research. *Journal of Information Technology Education*, 6 (2), pp. 1-20.

- Smith, A.J., Thurkettle, M.A. and de la Cruz, F.A. (2004) Use of intuition by nursing students: instrument development and testing. *Journal of Advanced Nursing*, 47(6), pp. 614-622.
- Speer, A.J., Solomon, D.J. and Ainsworth, M.A. (1996) An innovative evaluation method in an internal medicine clerkship. *Academic Medicine*, 71(1), pp. S76-8.
- Stake, R.E. (1977) The countenance of educational evaluation, in: Bellack, A.A. and Kliebard, H.M. (Eds.) *Curriculum and evaluation*. Berkeley, CA: McCutchan, pp. 372-390.
- Stake, R.E. (1983) Program evaluation, particularly responsive evaluation, in: Madaus, G.F., Scriven, M. and Stufflebeam, D.L. (eds.) *Evaluation models*, Dordrecht: Springer, pp. 287-310.
- Stange, K.C., Miller, W.L., Crabtree, B.F., O'Conner, P.J. and Zyzanski, S.J. (1994). Multimethod research: Approaches to integrating qualitative and quantitative methods. *Archives of General Internal Medicine*, 9(5), pp. 278-282.
- Steinaker, N.W. and Bell, M.R. (1979) *The Experiential Taxonomy: a new approach to teaching and learning*. London: Academic Press.
- Strauss, C., Taylor, B.L., Gu, J., Kuyken, W., Baer, R., Jones, F. and Cavanagh, K. (2016) What is compassion and how can we measure it? A review of definitions and measures. *Clinical psychology review*, 47(July), pp. 15-27.
- Stuart, C.C. (2009) *Assessment, supervision and support in clinical practice: A guide for Nurses and Midwives and other health professionals*. 2nd edn. London: Churchill Livingstone.
- Stuart, C.C. (2013) *Mentoring, Learning and Assessment in Clinical Practice: A Guide for Nurses, Midwives and Other Health Professionals*. London: Elsevier Health Sciences.
- Sumison, T. (1998) The Delphi Technique: an adaptive research tool. *The British Journal of Occupational Therapy*, 61(4), pp. 153-156.
- Tanicala, M.L., Scheffer, B.K. and Roberts, M.S. (2011) Defining pass/fail nursing student clinical behaviors phase I: Moving toward a culture of safety, *Nursing Education Perspectives*, 32 (3), pp. 155-161.
- Tanner, C.A. (2006) Thinking like a nurse: A research-based model of clinical judgement in nursing. *Journal of Nursing Education*, 45(6), pp. 204-211
- Tashakkori, A. and Teddlie, C. (1998). *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. London: Sage.
- Tashakkori, A. and Teddlie, C. (2010) *Sage handbook of mixed methods in social and behavioral research*. London: Sage.
- Teddlie, C., and Tashakkori, A. (2003) Major issues and controversies in the use of mixed methods in the social and behavioral sciences, in: Tashakkori, A. and Teddlie, C. (Eds.) *Handbook of mixed methods in social behavioral research*. Thousand Oaks, CA: Sage, pp. 3-50.
- Teddlie, C. and Yu, F. (2007) Mixed methods sampling: A typology with examples. *Journal of mixed methods research*, 1(1), pp. 77-100.
- Terry, K. (2013) The competency landscape: a critical realist exploration of the ways nurses understand and utilise competency standards, Doctoral dissertation, University of Tasmania.
- Terry, K., Stirling, C., Bull, R. and Fassett, D. (2017) An overview of the ways nurses understand and utilise the existing Australian Competency Standards for Registered Nurses. *Collegian*, 24(2), pp. 109-116.

- Thabane, L. and Akhtar-Danesh, N. (2008) Guidelines for reporting descriptive statistics in health research. *Nurse researcher*, 15(2), pp. 72-81
- Timmerman, B.E.C., Strickland, D.C., Johnson, R.L. and Payne, J.R. (2011) Development of a 'universal' rubric for assessing undergraduates' scientific reasoning skills using scientific writing. *Assessment and Evaluation in Higher Education*, 36(5), pp. 509-547.
- Tonesk, X. and Buchanan, R.G. (1987) An AAMC pilot study by 10 medical schools of clinical evaluation of students. *Journal of Medical Education*, 62(9), pp. 707-718.
- Torraco, R.J. (2005) Writing integrative literature reviews: Guidelines and examples. *Human resource development review*, 4(3), pp. 356-367.
- Trede, F., and M. Smith (2012) Challenges of Assessment in Practice-based Education, in: Higgs, J., Barnett, R., Billett, S., Hutchings, M. and Trede, F. (Eds.) *Practice-based Education. Perspectives and Strategies*. Rotterdam: Sense, pp. 187-198
- United Kingdom Central Council for Nursing Midwifery and Health Visiting (1986) *Project 2000: a new preparation for practice*. UKCC: London.
- United Kingdom Central Council for Nursing Midwifery and Health Visiting (1999) *Fitness for Practice: The UKCC commission for nursing and midwifery education*. UKCC: London.
- United Kingdom Central Council for Nursing, Midwifery, and Health Visiting (2001) *Fitness for Practice and Purpose: The Report of the UKCC's Post-Commission Development Group*. UKCC: London
- Ulrich, W. (2007) Philosophy for professionals: Towards critical pragmatism. *The Journal of the Operational Research Society*, 58(8), pp. 1109-1113.
- Valloze, J. (2009) Competence: A concept analysis. *Teaching and Learning in Nursing*, 4(4), pp. 115-118.
- Von der Gracht, H. (2008) The Delphi technique for futures research, in: Von der Gracht, H. (Ed) *The Future of Logistics: Scenarios for 2025*. Wiesbaden, Germany: GWV Fachverlage GmbH, pp.21-68.
- Watson, N.A. (1999) Mentoring today—the students' views. An investigative case study of pre-registration nursing students' experiences and perceptions of mentoring in one theory/practice module of the Common Foundation Programme on a Project 2000 course. *Journal of Advanced Nursing*, 29(1), pp. 254-262.
- Watson, R., Stimpson, A., Topping, A., and Porock, D. (2002) Clinical competence assessment in nursing: a systematic review of the literature. *Journal of advanced nursing*, 39(5), pp. 421-431.
- Weaver, K. and Olson, J.K. (2006) Understanding paradigms used for nursing research. *Journal of advanced nursing*, 53(4), pp. 459-469.
- Weaver, T. (1982) Policy options for post-tertiary education. *Higher Education Review*, 14(2), pp. 9-18.
- Webb, C. (1999) Analysing qualitative data: computerized and other approaches. *Journal of advanced nursing*, 29(2), pp. 323-330.
- Webb, C. and Shakespeare, P. (2008) Judgements about mentoring relationships in nurse education. *Nurse Education Today*, 28(5), pp. 563-571.
- Weber R.P. (1990) *Basic Content Analysis*. CA: Sage.
- Weiss, C.H. (1972) *Evaluation research*. Englewood Cliffs, NJ: Prentice-Hall.
- Weiss, C. H. (1998) *Evaluation: Methods for Studying Programs and Policies*, 2nd ed. Upper Saddle River, NJ: Prentice Hall.

- White, R. (1985) *The Effects of the NHS on the Nursing Profession 1948–1961*. King's Fund, London.
- White, S.K. (2004). The very idea of a critical social science: a pragmatist turn. In F. Rush (ed.), *The Cambridge Companion to Critical Theory*, UK: Cambridge University Press, pp. 310-335.
- Whiteford, G. (2007) Autonomy, accountability and professional practice: contemporary issues and challenges. *New Zealand Journal of Occupational Therapy*, 54 (1), pp. 11-14.
- Whittemore, R. and Knafl, K. (2005) The integrative review: updated methodology. *Journal of advanced nursing*, 52(5), pp. 546-553.
- Wiggins, G. (1989) A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70(9), pp. 703-713.
- Wilkes, Z. (2006) The student-mentor relationship: a review of the literature. *Nursing Standard*, 20 (37), pp. 42-47.
- Williams, P.L. and Webb, C. (1994) The Delphi technique: a methodological discussion. *Journal of advanced nursing*, 19(1), pp. 180-186.
- Williamson, G.R. and Webb, C. (2001) Supporting students in practice. *Journal of Clinical Nursing*, 10(2), pp. 284-292.
- Willis, D.O. (2009) Using competencies to improve dental practice management education. *Journal of Dental Education*, 73 (10), pp. 1144-1152.
- Willis Commission (2012) *Quality with compassion: the future of nursing education*. Royal College of Nursing, London.
- Willis, J. (2007) Assessment for learning—Why the theory needs the practice. *International Journal of Pedagogies and Learning*, 3(2), pp. 52-59.
- Wolfe, J. (2012) Does Pragmatism Have A Theory of Power? *European Journal of Pragmatism and American Philosophy*, 4(1), pp. 120-37.
- Wolf, K. and Stevens, E. (2007) The Role of Rubrics in Advancing and Assessing Student Learning. *Journal of Effective Teaching*, 7(1), pp. 3-14.
- Wood, M.J. and Ross-Kerr, J.C. (2011) *Basic steps in planning nursing research: From question to proposal*. 7th edn. London: Jones and Bartlett.
- Wu, X.V., Heng, M.A., Wang, W. (2015) Nursing students' experiences with the use of authentic assessment rubric and case approach in the clinical laboratories. *Nurse Education Today*, 35 (4), pp. 549-555.
- Yanhua, C. and Watson, R. (2011) A review of clinical competence assessment in nursing. *Nurse Education Today*, 31(8), pp. 832-836.
- Yorke, M. (2005) *Issues in the assessment of practice-based professional learning*. Milton Keynes: Open University.

Appendices

Appendix 1: Copy of the professional values competency statements in the PLPAD.

Appendix 2: The post-it notes used in the scoping exercise.

Appendix 3: Databases used to conduct the search in the integrative review.

Appendix 4: MMAT (Pace *et al.*, 2012) used in the integrative review.

Appendix 5: Round one e-Delphi questionnaire (group one).

Appendix 6: Round two e-Delphi questionnaire.

Appendix 7: University Research Ethical Committee approval for the Delphi study.

Appendix 8: University Ethical for the questionnaire developed after round one as a supplement.

Appendix 9: Invitation email sent to potential participants in round one of the Delphi.

Appendix 10: Participants information sheet sent to potential participants in the Delphi.

Appendix 11: A reminder email to potential participants in round one of the Delphi after two weeks.

Appendix 12: A final reminder email to potential participants in round one after three weeks.

Appendix 13: Invitation email sent to potential participants in round two of the Delphi.

Appendix 14: A reminder email to potential participants in round two of the Delphi after 1 week.

Appendix 15: A final reminder email to potential participants in round two after two weeks.

Appendix 16: The first draft version of the scoring rubric after phase one.

Appendix 17: Summary of the free-text comments in round two and the actions taken.

Appendix 18: Mentors questionnaire in phase two.

Appendix 19: Students questionnaire in phase two.

Appendix 20: HEI ethical approval for phase two.

Appendix 21: Permission from Head of Department at the HEI to access students.

Appendix 22: NHS REC online tool.

Appendix 23: Approvals from individual Trusts to access mentors.

Appendix 24: Indemnity letter from the HEI.

Appendix 25: Participants information sheet for phase two (students).

Appendix 26: Participants information sheet for phase two (mentors).

Appendix 27: Informed consent form for participating in phase two.

Appendix 28: A reminder email to participants in phase two after one week.

Appendix 29: A final email reminder to participants in phase two after two weeks.

Appendix 30: Key messages from the qualitative responses in phase two.

Appendix 1: Copy of the professional values competency statements in the PLPAD

Practice Assessment Document

Professional Values in Practice

You are required to demonstrate high standards of professional conduct at all times during your placements. As a student you should work within legal frameworks, and be able to articulate the underpinning values of the *NMC Code of professional conduct: standards for conduct, performance and ethics (2010)*. Professional values expectations are reflected in the statements below.

A = Achieved, N A = Not Achieved (Refer to Grade Descriptors on Page 5)

Professional attitude, behaviour and responsibility				
	Mid-Point	Final	Evidence/Comments	Final Sign/Date
1. The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and protection from harm.				
2. The student is non-judgemental, respectful and courteous at all times when interacting with patients/service users and all colleagues.				
3. The student maintains an appropriate professional attitude regarding punctuality and communicates appropriately if unable to attend placement.				
4. The student's personal presentation and dress code is in accordance with the organisation's uniform policy.				
5. The student acts as a role model in promoting a professional image.				
6. The student is proactive in promoting and maintaining the person's privacy and dignity.				
7. The student demonstrates openness, trustworthiness and integrity				

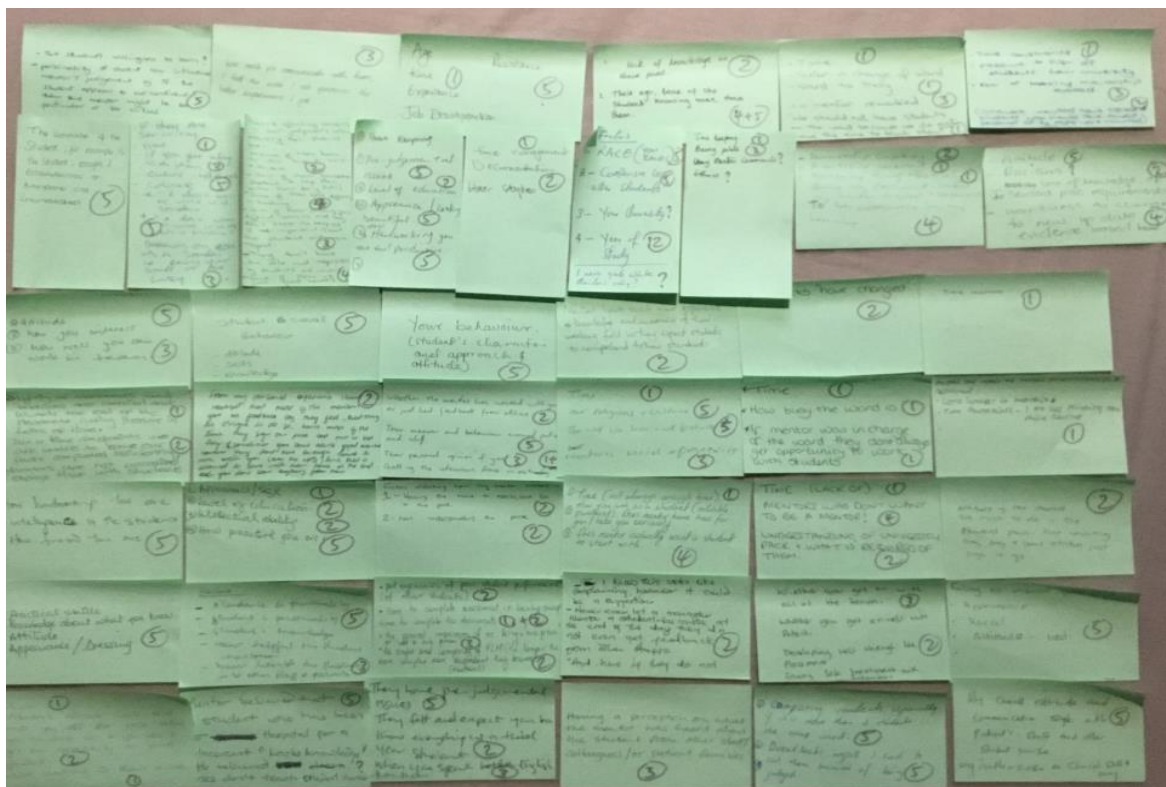
At the Mid-Point Interview, the Professional Values assessment is signed and dated at the end of the Mid-Point Interview. At the Final Interview signed and dated here.

Practice Assessment Document

	Mid-Point	Final	Evidence/Comments	Final Sign/Date
8. The student makes consistent effort to engage in and reflect on the requisite standards of evidence based care and learning to enhance care and their own professional development.				
Safe and compassionate care				
9. The student is attentive, kind, compassionate and sensitive to the needs of others.				
10. The student maintains consistent person-centred practice and empowers people to meet their own needs and make choices.				
11. The student reports any concerns to the appropriate professional member of staff when appropriate e.g. safeguarding.				
12. The student demonstrates the potential to lead and work autonomously and to listen and seek clarity where appropriate.				
13. The student is self-aware and self-confident and is able to work within the limitations of own knowledge, skills and professional boundaries and to take appropriate action.				
14. The student manages appropriate and constructive relationships with the multidisciplinary team, patients/service users, families and other carers, with the intent of building professional, caring relationships ensuring that decisions about care are shared.				

By the end of placement, not achieved must trigger an Action Plan at the time of assessment and should be documented. The Action Plan template can be found on Page 62

From the PLPAD (Version 2) Adult Nursing Part 3 BSc (2015) with permission from the Chair of Pan London Practice Learning Group.

[illegible]

Categories: 1. Organisational pressure. 2. Processes. 3. Relationships and emotions. 4. Mentors' aptitude. 5. Decision-making

Appendix 3: Databases used to conduct the search in the integrative review.

AMED (The Allied and Complementary Medicine Database): Contains bibliographic records from hundreds of journals.

BNI (British Nursing Index): Is a United Kingdom nursing and midwifery database, much smaller than CINAHL but does contain educational material (largely British) not found on other large databases.

CINAHL (Cumulative Index to Nursing and Allied Health Literature): Is the world's largest and most comprehensive database for nursing and the professions allied to medicine with over 100.000 educationally relevant citations.

Cochrane Library: is a collection of systematic reviews and meta-analyses in medicine and other healthcare specialties.

DARE (Database of Abstracts of Reviews of Effects): contains details of systematic reviews that evaluate the effects of healthcare interventions and the delivery and organisation of health services.

EBSCOhost (Elton B. Stephens Co.): is one of the largest privately held companies that provides a range of library database services.

EMBASE (Excerpta Medica database): Owned by Elsevier Science in Netherland and contains over 43.000 citations indexed as medical education

ERIC (Education Resource Information Centre): The Education Resource Information Centre is the world's largest education database although the emphasis is on school education, there are over 17.000 citations related to medical education.

ERC (Education Research Complete) is much smaller than ERIC but contains citations that are directly relevant to medical education.

ETHOS (Electronic Theses Online Service): developed by several UK Higher Education Institutions and the British Library. The website allows free access to over 250,000 doctoral theses records the full text of UK doctoral theses that have been digitised.

Google Scholar: Provides a simple way to broadly search for scholarly literature across many disciplines and sources from one place.

Joanna Briggs Institute: Evidence Based Practice database of systematic reviews complements those found in the Cochrane Library and develops evidence in various formats for nursing, allied health and medical professionals.

Medline (Medical Literature Analysis and Retrieval): The National Library of Medicine in the USA. It contains over 73.000 citations indexed as medical education and over 300,000 citations that are considered educationally relevant.

PsycINFO: Concentrates on psychiatric and psychological content, there are over 4,000 records indexed as medical education with well over 100,000 concerning education in a broader context.

ScienceDirect: Is a website which provides access to a large database of scientific and medical research. It hosts over 12 million pieces of content from 3,500 academic journals.

Appendix 4: The mixed method assessment tool

Types of mixed methods study components or primary studies	Methodological quality criteria (see tutorial for definitions and examples)	Responses			
		Yes	No	Can't tell	Comments
Screening questions (for all types)	Are there clear qualitative and quantitative research questions (or objectives*), or a clear mixed methods question (or objective*)?				
	Do the collected data allow address the research question (objective)? E.g., consider whether the follow-up period is long enough for the outcome to occur (for longitudinal studies or study components).				
	<i>Further appraisal may be not feasible or appropriate when the answer is 'No' or 'Can't tell' to one or both screening questions.</i>				
1. Qualitative	1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?				
	1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?				
	1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?				
	1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?				
2. Quantitative randomized controlled (trials)	2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?				
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)?				
	2.3. Are there complete outcome data (80% or above)?				
	2.4. Is there low withdrawal/drop-out (below 20%)?				
3. Quantitative non-randomised	3.1. Are participants (organizations) recruited in a way that minimizes selection bias?				
	3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?				
	3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?				
	3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?				
4. Quantitative descriptive	4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?				
	4.2. Is the sample representative of the population understudy?				
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)?				
	4.4. Is there an acceptable response rate (60% or above)?				
5. Mixed methods	5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and				

	quantitative aspects of the mixed methods question (or objective)?				
	5.2. Is the integration of qualitative and quantitative data (or results*) relevant to address the research question (objective)?				
	5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results*) in a triangulation design?				
	<i>Criteria for the qualitative component (1.1 to 1.4), and appropriate criteria for the quantitative component (2.1 to 2.4, or 3.1 to 3.4, or 4.1 to 4.4), must be also applied.</i>				

*These two items are not considered as double-barrelled items since in mixed methods research, (1) there may be research questions (quantitative research) or research objectives (qualitative research), and (2) data may be integrated, and/or qualitative findings and quantitative results can be integrated.

Appendix 5: Round one e-Delphi questionnaire (group one).

London South Bank
University

Delphi study group 1

Page 1: Welcome to the Delphi survey

Thank you for participating in the survey. If you have any questions related to the study, please do not hesitate to contact me.

Clicking on the link below indicates that you have read the participant information sheet and you voluntarily agree to participate in the survey.

Ibraheim Almalkawi

Email: [REDACTED]

Phone: [REDACTED]

Page 2: Some questions about you

You are completing this questionnaire as a

- ☐ Mentor
- ☐ Student
- ☐ Practice facilitator
- ☐ Service user
- ☐ Academic
- ☐ Other

If you selected Other, please specify: *Optional*

Page 3: Complete this page only if you are a mentor. If not please go to the next page?

As a qualified mentor, how many years have you been assessing students

- ☐ 0-5 years
- ☐ 6-10 years
- ☐ more than 10 years

In the last year, you have been the main mentor for how many students

- ☐ None
- ☐ 1 student
- ☐ 2 students
- ☐ 3 or more

Are you a sign off mentor

- ☐ Yes
- ☐ No

Page 4: Complete this page only if you are a service user. If not please go to the next page?

As a service user, in what capacity are you involved with student nurses

- ☐ Have been cared for by student nurses
- ☐ Assessed student nurses
- ☐ Curriculum development for student nurses
- ☐ Teaching student nurses
- ☐ Recruiting student nurses
- ☐ Other

If you selected Other, please specify:

Page 5: Professional values in practice 1

In your interpretation of the statement below, describe what the expected performance looks like making distinction across the 3 levels.

[More info](#)

	Don't want to see	Expect to see	Love to see
The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and protection from harm.			

Page 6: Professional values in practice 2

In your interpretation of the statement below, describe what the expected performance looks like making distinction across the 3 levels.

[More info](#)

	Don't want to see	Expect to see	Love to see
The student is nonjudgmental, respectful and courteous at all times when interacting with patients/service users and all colleagues.			

Page 7: Professional values in practice 3

In your interpretation of the statement below, describe what the expected performance looks like making distinction across the 3 levels.

 [More info](#)

	Don't want to see	Expect to see	Love to see
The student maintains an appropriate professional attitude regarding punctuality and communicates appropriately if unable to attend placement.			

Page 8: Professional values in practice 4

In your interpretation of the statement above, describe what the expected performance looks like making distinction across the 3 levels.

 [More info](#)

	Don't want to see	Expect to see	Love to see
The student's personal presentation and dress code is in accordance with the organisation's uniform policy.			

Page 9: Thank you very much for your contribution to this research. I will be in touch with you again soon to share the summary statements of all the feedback in a second questionnaire.

- Group 2 questionnaire was identical only differ in the competency statements page 5-8 (All competency statement can be seen in appendix 5).

Appendix 6: Round two e-Delphi questionnaire.

Round 2 Delphi survey

Page 1: Welcome to round 2 of the Delphi survey

Thank you for participating in the survey. If you have any questions related to the study, please do not hesitate to contact me.

Clicking on "Next" below indicates that you voluntarily agree to participate in the survey

Ibraheim Almalkawi

Email

Phone

Page 2: Some questions about you

You are completing this questionnaire as a

- ☐ Mentor
- ☐ Student
- ☐ Practice facilitator
- ☐ Service user
- ☐ Academic
- ☐ Other

If you selected Other, please specify: *Optional*

Page 3: Complete this page only if you are a mentor. if not please go to the next page.

As a qualified mentor how many years have you been assessing students?

- ☐ 0-5 years
- ☐ 6-10 years
- ☐ more than 10 years

In the last year, you have been the main mentor for how many students?

- ☐ None
- ☐ 1 student
- ☐ 2 students
- ☐ 3 or more

Are you a sign off mentor?

- ☐ Yes
- ☐ No

Page 4: Complete this page only if you are a service user. If not please go to the next page.

As a service user in what capacity are you involved with student nurses?

- ☐ Have been cared for by student nurses
- ☐ Assessed student nurses
- ☐ Curriculum development for student nurses
- ☐ Teaching student nurses
- ☐ Recruiting student nurses
- ☐ Other

If you selected Other, please specify:

Page 5: Professional value 1

The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and self-harm.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: Discloses details about patients, colleagues or clinical area in public including social media. Discusses information with staff not involved in the care or within earshot of others. Does not appreciate limits to confidentiality and need to raise concern for safeguarding purposes. Lacks attention regarding the protection and disposal of written information.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Consistently maintains confidentiality in their day to day practice showing correct processes for discussing, accessing, sharing, storing and disposing of sensitive data in all formats. Understands the principles of public protection and individual safeguarding in situations where confidentiality should be breached and able to illustrate the appropriate processes to disclose information.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Recognises the boundaries of confidentiality in all forms of communications and is skilful in balancing the sensitivity of maintaining confidentiality with the needs of distressed relatives to seek information. Understands the tensions that occur in practice by continuing to challenge others when they fail to comply with the confidentiality code yet champions the patient's rights to harm-free care.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comments?

Page 6: Professional value 2

The student is nonjudgmental, respectful and courteous at all times when interacting with patients/service users and all colleagues.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: Judges others based on their differences, e.g. culture, religion, race, gender, sexuality or their own values and beliefs. Being rude or aggressive to others or ignores their needs. Engages in inappropriate conversations, e.g. gossip, arguments or derogatory comments.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Treats others as individuals and with courtesy irrespective of their diversity while respecting their differences. Non-opinionated, unbiased, and acts professionally in a way that shows courtesy and politeness, both in language and manner.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Seeks to understand and is sensitive to the cultural and religious principles of patients and relatives, and advocates upholding such values on their behalf. Recognises own personal perspectives that may influence how they provide care. Is able to identify and reflect on situations where a conflict of values may arise and proactively seeks to address this.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 7: Professional value 3

The student maintains appropriate professional attitude regarding punctuality and communicates appropriately if unable to attend placement.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: Failure to adhere to start, finish and break times. Not reporting lateness or non attendance. Changes shifts without good reason and at short notice. Lacks understanding of the importance of punctuality on teamwork and patient care.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Consistently punctual. Adheres to sickness and absence policies and communicates all attendance issues to placement and university in a professional and timely manner and is proactive in planning and communicating their return following a period of absence. Makes rostering requests in a timely and considerate manner to facilitate any required planned absence.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Illustrates a positive attitude and willingness to gain as much as possible from the placement, (e.g. arriving early for shifts, planning learning opportunities). Demonstrates a good work ethic and is conscious of the impact of lateness/absence on patient care, the team they work with and the completion of the course they are studying. Is proactive in preparing for the allocated duty and is considerate to the placement needs before their own convenience.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 8: Professional value 4

The student's personal presentation and dress code is in accordance with the organisation's uniform policy.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: The uniform is dirty/creased or incorrect for the organisation. Name badge/ID not displayed or wears uniform outside placement without a reason. Scruffy or unprofessional appearance (e.g. hair, nails, makeup, jewellery, perfume) or poor hygiene. Lacks knowledge about the organisation's uniform policy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Consistently looks presentable and clean. Wears the correct uniform for each shift which is clean, well pressed and displaying ID clearly. Able to explain the importance of presentation, dress code, cleanliness and their association to professionalism and patient safety.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Presents themselves in a professional manner that demonstrates pride in wearing the uniform. Acts as a role model to promote and encourage others to appreciate the importance of professional appearance. Challenges others when their personal presentation or hygiene standards may affect patients' safety or confidence.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 9: Professional value 5

The student acts as a role model in promoting a professional image.

[+ More info](#)

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: Exhibits behaviour, appearance and/or body language that reflects lack of interest in nursing/healthcare or understanding of how inappropriate behaviour and/or action can impact on the profession and/or the organisation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Sees self as representative of the profession and organisation. Takes pride and sets an example of acting in a way to uphold and promote professional values and passion for nursing in all interactions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Challenges others who may discredit the profession/organisation. Contributes to enhance public perceptions of nursing and speaks up when nursing is inaccurately portrayed. Engages in and advocates for action on health issues of public importance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 10: Professional value 6

The student is proactive in promoting and maintaining the individual's privacy and dignity.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: Unnecessarily exposes patients during care (e.g. not closing curtains, doors, windows) or disregards privacy measures (e.g. enters or allows others entry without permission). Not taking steps to enable patients (or carers) to carry out own personal needs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Demonstrates knowledge and shows consideration concerning the importance of privacy and dignity on the patient's quality of care and always takes steps to ensure all measures are taken without being prompted. Fulfils their role in promoting privacy and dignity by correcting or escalating failings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Identifies challenges to privacy and dignity and shapes their own care and that of others to ensure this is maintained. Takes time to find out more about their patient's needs to appreciate personal cultural and religious variations to privacy in order to provide individualised care.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 11: Professional value 7

The student demonstrates openness, trustworthiness and integrity.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: Not trusted to communicate or document accurate account of events or keep promises. Not being open in admitting mistakes made or blames others. Failure to admit own limitations in experience, knowledge or skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Truthful, delivers the care that has been committed to, keeps clear and timely records and communicate/ handover care or events accurately. Trusted to carryout patient care and honest about own limitations. Openly discusses mistakes and accepts responsibility for any short comings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Exhibits professional conduct that reflects awareness of being a role model in representing the profession in all that they say or do, and in giving accurate accounts and documentation of events. Takes responsibility for their own actions and reports mistakes swiftly and openly to minimise harm and is an advocate for such behaviour.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 12: Professional value 8

The student makes consistent effort to engage in and reflect on the requisite standards of evidence based care and learning to enhance care and their own professional development.

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Don't want to see: No consideration for evidence/ rationale/ research to support practice. Passive learner, does not seek out information that will enhance care or own learning. Lacks interest or recognition of responsibility for learning and development. No efforts to reflect on personal and professional development.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expect to see: Has a questioning approach to care and challenges poor practice. Is keen to utilise research and local/national policy to underpin care delivery to enhance patient care and safety. Reflects on their practice to identify personal learning needs, and engages in shared learning/practice development and seeks out learning opportunities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Love to see: Uses research databases/journals and can critically appraise the quality of evidence before applying it to practice. Identifies and challenges practice that lacks an evidence base and works with others to address this. Up to date with local/national agendas or directives and acts as a resource to colleagues.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Any comment?

Page 13

Please rank the following level descriptors in the order of your preference to use in the final version of the rubric?

	Strongly Favour	Favour	Neutral	Dislike	Strongly Dislike
Don't want to see - Expect to see - Love to see	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Developing - Competent - Exemplary	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Does not meet expectations - Meets expectations - Exceeds expectations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Not achieved - Achieved - Merit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1- 2 - 3 (Numerical where 2 is needed for a pass	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Page 14: Thank you very much for your contribution to this research.

Appendix 7: University Research Ethical Committee approval for the Delphi study.

Direct line: 020 [REDACTED]
E-mail: [REDACTED]@[REDACTED].ac.uk
Ref: UREC 1520

Ibraheim Almalkawi

Wednesday 26 August 2015

Dear Ibraheim

RE: Developing a rubric for mentor assessment of pre-registration nursing students: Stage 1

Thank you for submitting this proposal and for your response to the reviewers' comments.

I am pleased to inform you that Full Chair's Approval has been given by Professor [REDACTED] on behalf of the University Research Ethics Committee.

I wish you every success with your research.

Yours sincerely,

[REDACTED]
[REDACTED]
Secretary, [REDACTED] Research Ethics Committee

cc:

Prof [REDACTED], Chair, [REDACTED] Research Ethics Committee

Appendix 8: University Research Ethical Committee approval for the questionnaire developed after round one as a supplement.

[Reply](#) [Reply All](#) [Forward](#)

RE: UREC 1520

Governance : Administration Of Research Ethics Committee

To: Almalkawi, Ibraheim

Dear Ibraheim

I am pleased to confirm that Vice Chair, [REDACTED], has looked at the round 2 Delphi questionnaire in conjunction to this and I agree that it can be approved as an amendment to the original application (UREC 1520) in line with Delphi survey methodology.

Please let me know if you need an approval letter for this.

Kind regards

[REDACTED]

Administrative Assistant
University Research Ethics Committee
[REDACTED] University
[REDACTED] Office
[REDACTED] .uk

From: Almalkawi, Ibraheim
Sent: 12 November 2015 09:42
To: Governance : Administration Of Research Ethics Committee
Cc: [REDACTED]
Subject: RE: UREC 1520

Dear [REDACTED]

Please find attached the questionnaire for round 2 Delphi to supplement the ethics application UREC 1520. I was instructed to supply it once it has been prepared.

Regards

Ibraheim

Appendix 9: Invitation email sent to potential participants in round one of the Delphi.

Dear Sir/Madam

I am writing to invite you to participate in a study that is part of a doctoral degree at xxxxx University. The aim is to identify how different stakeholders interpret the competency statements of the practice assessment document for nurse training. Ethical approval for this study was granted by xxxx university on the 26th August 2015 (UREC 1520).

This study will use a technique known as 'Delphi' which involves identifying a group of stakeholders with relevant input about a particular field and seek their opinions about it. This means there are no right or wrong answers to the questions. Participants will be asked to complete 2 or possibly 3 online questionnaires (known as rounds). The amount of time required to complete each questionnaire will vary with each panellist and range from 30-45 minutes for round 1 and 15-20 minutes for subsequent rounds. I have attached an information sheet which gives you further information about taking part.

I would be very grateful if you give permission to participate in this study and I hope you will find the process interesting in informing your understanding of how your own opinion is positioned among others. More importantly, inconsistencies in the way mentors interpret competency statements are well established in research studies, and you will be contributing towards designing a tool to help mentors and students reach a common language in defining what is expected to be achieved when assessing competencies in practice placements. It is important that you understand that your participation is voluntary and you will remain anonymous throughout the study.

I sincerely hope that you will agree to participate. If you are interested in taking part, please click on the link below to access the questionnaire.

Thank you for your time and any help you may be able to offer to this study. If you have any questions please email or call me.


[Click here to begin the survey](#)

Yours sincerely

Ibraheim Almalkawi (PhD Student)

School of Health and Social Care, xxxxxx University, Email: xxxxxx@xxxxx.ac.uk, Phone: +44(0)20 7815 xxxx

Appendix 10: Participants information sheet sent to potential participants in the Delphi

	UREC No. 1520
Participant Information Sheet	
Title: Developing a rubric (marking grid) to enhance mentors' assessment of student nurses' practice performance	
<p>You are being invited to take part in a Delphi consensus study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish. Please contact me if you need further clarification or would like more information. Take time to decide whether or not you wish to take part.</p>	
<p>What is a Delphi study? The Delphi technique seeks to obtain consensus of opinions from people with relevant input about a topic (termed panel members) through a series of questionnaires. As part of the process, responses are fed back to the members who are then given an opportunity to respond again to the emerging data. The Delphi is therefore a multi-stage process designed to combine opinion into group consensus.</p>	
<p>What is the purpose of the study? The literature suggests that mentors are not always able to make effective assessments about student nurses due to the way competency statements are interpreted. These findings emphasise the need to define a transparent and common language that is helpful for assessors to identify what is expected of a student and for the student to identify and understand what they are expected to achieve. The aim of this study is to explore whether the panel members could reach a consensus on a common interpretation of the current professional values competency statements including describing different levels of achievements.</p>	
<p>Why have I been invited to take part? As someone with relevant knowledge in this field, you are invited to participate in this study to interpret the competency statements and describe associated performance levels. The aim is to recruit 50 members including practice assessors, students, informed nursing academics, strategic or policy shapers and service users. If you agree to take part, you will be sent an email with a link directing you to the survey website. Accessing the website and completing the survey indicates that you consent to voluntarily participate. You are free to withdraw from the study at anytime without the need for explanation and you will not be disadvantaged.</p>	
<p>What will I be asked to do if I take part? If you are willing to participate as a Delphi panel member, this would involve answering 2-3 questionnaires (each questionnaire is termed 'round') related to competency statements using an online survey. It is envisaged that this should take approximately 30-45 minutes for the first round and 15-20 minutes for subsequent rounds. In the first round you will be provided with free text boxes to provide your interpretations of the competency statements. Future rounds will be on average 4-6 weeks apart, where you would receive a summary of the group's interpretation formed into statements and you will be asked to rate your level of agreement or disagreement with the statements using Likert scales ranging from strongly agree to strongly disagree. This process would</p>	
Page 1 of 2	

continue until a group consensus is achieved or three rounds are completed. In order to allow timely completion of the study your response to each round would be required within 2 weeks.

What are the possible benefits and risks of this study

It is not anticipated that you will be disadvantaged or suffer any risk from this study. However, if you experience any distress please remember that you can stop completing the survey and contact me for support.

It is unlikely that you will gain any personal benefit from participating in this research, but your input is valuable in attempting to improve the accuracy of assessment of student nurses in the practice setting and to minimise the risk of inaccurate passing or failing of students.

Will my talking part in this study be kept confidential?

Data will be collected through a secured website where only invited members will be allowed access. All information received from you will be collated anonymously, handled in a confidential manner and stored in a locked filing cabinet and on a password protected computer in an environment locked when not occupied. All responses received in the study will be strictly confidential, and your identity will not be disclosed. Direct quotes to free-text answers may be used in the study report, but will not be traceable back to you. Should any issue of concern be revealed in the survey, I will discuss the issue with you and together decide whether to raise the concern.

Only I and my supervisors will have direct access to the information. Any reference to you will be coded. This information will be held for a maximum of 5 years after the PhD project is completed and then it will be destroyed.

Who has reviewed the study?

This study is being completed as part of a PhD degree at London South Bank University. It has been reviewed and ethically approved by the University Research Ethics Committee.

Who do I contact if for more information or if I have concerns?

If you wish for any further information or you have concerns about any aspect of this study or experience any distress you can stop completing the survey and feel free to speak me. If you have any complaints about the way you have been dealt with during the study or other concerns you can contact Dr. [REDACTED] at 020 [REDACTED] who is the Academic Supervisor for this study. Finally, if you remain unhappy and wish to complain formally, you can contact the Chair of the University Research Ethics Committee. Details can be obtained from the university website:

[https://\[REDACTED\].ac.uk/page/research-degrees-ethics](https://[REDACTED].ac.uk/page/research-degrees-ethics)

What do I do now?

Thank you for reading this information sheet and for considering taking part in this research. Please respond to this email within a week to let me know if you wish to take part in the study.

If you have any questions or concerns please do not hesitate to contact me.

Ibraheim Almalkawi
PhD Student

[REDACTED].ac.uk
020 [REDACTED]

Professor [REDACTED]
Director of studies

[REDACTED].ac.uk

Dr. [REDACTED]
Supervisor

[REDACTED].ac.uk
020 [REDACTED]

Appendix 11: A reminder email to potential participants in round one of the Delphi after two weeks

Dear Sir/Madam

I recently wrote inviting you to take part in a survey which seeks to identify how different stakeholders interpret the competency statements of the practice assessment document.

This survey is part of a PhD research project into the reported inconsistency in mentors' interpretations of clinical competencies in pre-registration nursing training aiming to articulate common language to differentiate between levels of performance. Therefore I am very interested in your views and your response is very important to this study to reach a collective interpretation of the statements. Hence I am writing to encourage you to complete the survey.

There are no right or wrong answers to the questions and I have attached the participant information sheet which gives you further information about taking part.

The views sought in this research covers all stakeholders including students, mentors, academics, practice educators and policy makers. All of the responses will be treated in the strictest confidence and survey will be stored in anonymised form.

If you have any queries, please feel free to contact me at xxxx@xxxx.ac.uk.

I realise the great pressures on your time and would like to thank you in advance for taking the trouble to read this e-mail and, I hope, to complete the survey.

[PLEASE FOLLOW THIS LINK TO TAKE THE SURVEY](#)

Yours sincerely
Ibraheim Almalkawi (PhD Student)
Phone: +44(0)20 xxxx xxxx

Appendix 12: A final reminder email to potential participants in round one of the Delphi after three weeks

Dear Sir/Madam

If you have already completed the questionnaire, thank you very much for doing this and please accept my apology for sending this reminder.

This is just a final prompt for those still wanting to do it that it's not too late!

Your participation in this study is very valuable; I appreciate you taking time from your busy schedule to contribute to this study.

[Click here to begin the survey](#)

Yours sincerely

Ibraheim Almalkawi (PhD Student)

Email: xxxxx@xxxx.ac.uk

Phone: +44(0)20 xxxx xxxx

Appendix 13: Invitation email sent to potential participants in round two of the Delphi

Dear Participant

Round 1 of the Delphi is now completed, thank you to all participants. The responses were extensive, thoughtful, and very helpful.

Please find below the link to round 2 which contains the analysed responses from round 1. When you click on the link you will be asked to assess and comment on the appropriateness of the statements. I would value your participation in this round even if you didn't complete round 1.

This round will take approximately 10 minutes and participation implies consent.

I would respectfully ask for a response time of 2 weeks (Wed, November 25th 2015). After one week, you will receive a reminder email to complete the survey if you haven't already done so.

It is not clear at this stage if there will be a third round, this will depend on the degree of consensus in round 2.

Thank you for your time and help, if you have any questions please email or call me.

[Click here to begin round 2](#)

Kind Regards

Ibraheim Almalkawi (PhD Student)
Email: xxxx@xxxx.ac.uk
Phone: +44(0)20 xxxx xxxx

Appendix 14: A reminder email to potential participants in round two of the Delphi after one week.

Dear Sir/Madam

I recently invited you to take part in the 2nd Delphi survey. Your response is very important to establish if there is a consensus among different stakeholders on the interpretation of the competency statements. Hence I am writing to encourage you to complete the survey. You can still complete this questionnaire even if you didn't do the first one.

This survey will take approximately 5-10 minutes.

If you have already completed the questionnaire, thank you very much for doing this and please accept my apology for sending this reminder.

[Please follow the link to complete the survey](#)

Yours sincerely

Ibraheim Almalkawi (PhD Student)

Email: xxxxxx@xxxx.ac.uk

Phone: +44(0)20 xxxx xxxx

Appendix 15: A final reminder email to potential participants in round two of the Delphi after two weeks

Dear Sir/Madam

If you have already completed the second questionnaire, thank you very much for doing this and please accept my apology for sending this reminder.

This is just a final prompt for those still wanting to do it that it's not too late!

This questionnaire will take around 5 minutes to complete.

Your participation in this study is very valuable; I appreciate you taking time from your busy schedule to contribute to this study.

[Please follow the link to complete the survey](#)

Yours sincerely

Ibraheim Almalkawi (PhD Student)

Email: xxxxxx@xxxx.ac.uk

Phone: +44(0)20 xxxx xxxx

Appendix 16: The first draft version of the scoring rubric after phase one.

Professional values	Don't want to see	Expect to see	Love to see
The student maintains confidentiality in accordance with the NMC code and recognises limits to confidentiality for example public interest and self-harm	Discloses details about patients, colleagues or clinical area in public including social media. Discusses information with staff not involved in the care or within earshot of others. Does not appreciate limits to confidentiality and need to raise concern for safeguarding purposes. Lacks attention regarding the protection and disposal of written information.	Consistently maintains confidentiality in their day to day practice showing correct processes for discussing, accessing, sharing, storing and disposing of sensitive data in all formats. Understands the principles of public protection and individual safeguarding in situations where confidentiality should be breached and able to illustrate the appropriate processes to disclose information.	Recognises the boundaries of confidentiality in all forms of communications and is skilful in balancing the sensitivity of maintaining confidentiality with the needs of distressed relatives to seek information. Understands the tensions that occur in practice by continuing to challenge others when they fail to comply with the confidentiality code yet champion the patient's rights to harm free care.
The student is non-judgmental, respectful and courteous at all times when interacting with patients/service users and all colleagues.	Judges others based on their differences, e.g. culture, religion, race, gender or their own values and beliefs. Being rude or aggressive to others or ignores their needs. Engages in inappropriate conversations, e.g. gossip, arguments or derogatory comments.	Treats others as individuals and with courtesy irrespective of their diversity while respecting their differences. Non-opinionated, unbiased, and acts professionally in a way that shows courtesy and politeness, both in language and manner.	Seeks to understand and embrace the cultural and religious principles of patients and relatives, and advocates upholding such values on their behalf. Recognises own personal perspectives that may influence how they provide care, and is able to reflect on situations that caused them to question their own action or that of others.
The student maintains appropriate professional attitude regarding punctuality and communicates appropriately if unable to attend placement.	Non-attendance or adherence to start, finish and break times. Not reporting lateness or non-attendance. Changes shifts without good reason and at short notice. Lacks understanding of the importance of punctuality on teamwork and patient care.	Adheres to sickness and absence policies. Consistently punctual throughout the placement allocation. Communicates all attendance issues to placement and university in a professional and timely manner and proactive in planning and communicating their return following a period of absence. Makes rostering requests in a timely and considerate manner to facilitate any required planned absence.	Illustrates positive attitude and willingness to gain as much as possible from the placement, (e.g. arriving early for all shifts, planning learning opportunities). Demonstrates a good work ethic and is conscious of the impact of lateness/absence on patient care, the team they work with and the completion of the course they are studying. Proactive in preparing for the allocated duty and is considerate to the placement needs before their own convenience.
The student's personal presentation and dress code is in accordance with the organisation's uniform policy.	The uniform is dirty/creased or incorrect for the organisation. Name badge/ID not displayed or wears uniform outside placement without a reason. Scruffy or unprofessional appearance (e.g. hair, nails, makeup, jewellery, perfume) or poor hygiene. Lacks knowledge about the organisation's uniform policy.	Consistently looks presentable and clean. Wears the correct uniform for each shift which is clean, well pressed and displaying ID clearly. Able to explain the importance of presentation and dress code and cleanliness and it's association to professionalism and patient safety.	Presents themselves in a professional manner that demonstrates pride in wearing the uniform. Acts as role model to promote and encourage others to appreciate the importance of professional appearance. Challenges others when their personal presentation or hygienic standards may affect patients' safety or confidence.

The student acts as a role model in promoting a professional image.	Exhibits behaviour, appearance and/or body language that reflects lack of interest in nursing/healthcare or understanding of how inappropriate behaviour and/or action can impact on the profession and/or the organisation	Sees self as representative of the profession and organisation. Takes pride and sets an example of acting in a way to uphold and promote professional values and passion for nursing in all interactions.	Challenges others who may discredit the profession/organisation. Contributes to enhance the public perceptions of nursing and speaks up when nursing is inaccurately portrayed. Engages in and advocates for action on public health matters to help nursing as profession to gain a position of influence.
The student is proactive in promoting and maintaining the individual's privacy and dignity.	Unnecessarily exposes patients during care (e.g. curtains, doors, windows) or disregards privacy measures (e.g. enters or allows others entry without permission). Not taking steps to enable patients (or carers) to carry out own personal needs.	Demonstrates knowledge and shows consideration concerning the importance of privacy and dignity on patient's quality of care and always takes steps to ensure all measures are taken without being prompted. Fulfils their role in promoting privacy & dignity by correcting or escalating failings.	Thoughtful about challenges to privacy and dignity and shapes their own care and that of others to ensure this is maintained. Takes time to find out more about their patient's needs to appreciate cultural and religious variations to privacy in order to provide individualised care.
The student demonstrates openness, trustworthiness and integrity.	Not perceived as trusted to communicate or document accurate account of events or keep promises. Not being open in admitting mistakes made or blames others. Failure to admit own limitations in experience, knowledge or skills.	Truthful, delivers the care that has been committed to, keeps clear and timely records and communicate/ handover care or events accurately. Trusted to carryout patient care and honest about limitations; openly discusses mistakes and accept responsibility for any short comings.	Exhibits professional conduct that reflects awareness of being a role model in representing the profession in all that they say or do, and in giving accurate accounts and documentation of events. Takes responsibility for their own actions and reports mistakes swiftly and openly to minimise harm and is an advocate for such behaviour.
The student makes consistent effort to engage in and reflect on the requisite standards of evidence based care and learning to enhance care and their own professional development.	No consideration for evidence/ rationale/ research to support practice. Passive learner, does not seek out information that will enhance care or own learning. Lacks interest or recognition of responsibility for learning and development. No efforts to reflect on personal and professional development.	Has a questioning and challenging approach to poor practice and keen to utilise research and local/national policy to establish evidence underpinning care delivery to enhance patient care and safety. Reflects on their practice to identify personal learning needs, and engages in shared learning/practice development and seeks out learning opportunities.	Able to research databases/journals and can critically appraise the quality of evidence before applying it to their practice. Appropriately challenges poor practice and provides alternative solutions or points to evidence base. Up to date with local/national agendas or directives and acts as a resource to colleagues. Reflects on the clinical areas and identifies practices that lacks an evidence base and works with others to address this.

Appendix 17: Summary of the free-text comments in round two and the actions taken.

	Comment	Group	Response to comments	Action
1	Possibly my interpretation but in the love to see, I find the last sentence too wordy starting "understands.... Maybe change to "is able to challenge others in practice when they fail to comply to confidentiality code and champions.....	Mentor 1	The first sentence is about "understand the tensions" of challenging others in practice vs. champions the patients' rights.	None
	Depends on which year the student is in	Academic 1	The rubric is designed for 3 rd year students and this was made clear in the information provided.	None
	Only agree with first statement as at times it may be appropriate to discuss a patient's condition/situation with others not directly involved in care (eg. safeguarding concerns)	Other (joint academic/clinical commissioner) 1	The other statements clearly discuss the obligation to share information in safeguarding situations	None
2	I am not keen on the wording "non opinionated" as the nurse will still have opinions but do not let them influence how they treat others. Consider rephrasing this statement	Mentor 1	"Non-opinionated" may be ambiguous and could be misinterpreted as not having an opinion.	"Non-opinionated" changed to "unprejudiced"
	Very clear guidance	Academic 2	None	None
	The first question is unclear - I strongly agree that I don't want to see these attitudes. The way the question is phrased may be misleading - the same applies to the previous question (the same comment to all the questions)	(Commissioner 1)	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	Love to see students who actively seeks to understand	Practice educator 1	The comment suggested adding the word "Actively".	The paragraph now reads "Actively seeks .."
	I have a problem with terminology. Love to see.	Practice educator 2	It was identified in question 14 that most respondents favoured Does not meet expectations - Meets expectations - Exceeds expectations	The ranking changed to Does not meet expectations - Meets expectations - Exceeds expectations
3	I would make your love to see the expect to see category and go higher on the love to see someone about helping others, flexibility when covering for staff	Academic 2	The competency statement as it is includes the suggestion in the comment.	None
	Same comment as before	(Commissioner 1)	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	Would expect this in the 'Expect to See' Demonstrates a good work ethic and is conscious of the impact of lateness/absence on patient care, the team they work with and the completion of the course they	Practice educator 3	The higher level expectations are appropriate as it matches what is expected of a qualified nurse and was supported by the majority. However, to expect	"Demonstrates a good work ethic" has been replaced by

	are studying. Is proactive in preparing for the allocated duty and is considerate to the placement needs before their own convenience.		more than demonstrating good work ethic at the higher level may be appropriate.	"Act as role model in promoting good work ethic".
	With the last comment, I don't think arrives early for shifts should be in there as it could actually be disruptive if lots of students were arriving early	Academic 3	This needs to be rewarded so it does not mean arriving early physically but to mean coming prepared.	The statement will now read: "coming prepared"
	The student is sometimes impacted upon by Trust / placement culture and is influenced by other staff behaviour	(Nurse Education Lead)	Although this is an important statement, but this is not a competence for the student to achieve	None
4	Scruffy is an ambiguous word - alter it Again I would make the love to see the expect to see and discuss dress code as an organisational issue that facilitate communication and create confidence in the institute	Academic 2	The competency statement as it is includes the suggestions in the comment.	Scruffy changed to unkempt
	Same comment as before	(Commissioner 1)	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	Would Expect to see Presents themselves in a professional manner that demonstrates pride in wearing the uniform. Acts as a role model to promote and encourage others to appreciate the importance of professional appearance	Practice educator 3	The higher level expectations are appropriate as it matches what is expected of a qualified nurse and was supported by the majority.	None
	The first statement: I think this needs expressing more clearly. ie what is it about nails (dirty? too long?). Makeup: could imply it shouldn't be worn at all but some moderate makeup is ok. Rather than 'lacks knowledge about the organisation's uniform policy' should it be ignores or pays no regard, as none of the students can say they don't know about it	Academic 3	The unacceptable appearance implies dirty or long nails and further clarifications are not required. Ignoring or pays no regard reads better.	The statement will read "ignores or pays no regard" instead of "lacks knowledge about.."
	This should be for all employees of the organisation, not just students (though the study is aimed at students)	Darzi Fellow 1	None	None
5	Again dependant on year	Academic 1	The rubric is designed for 3 rd year students and this was made clear in the information provided.	None
	Clearest set of values so far	Academic 2	None	None
	Same comment as before	(Commissioner 1)	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	I feel nurses can be professional without having to have a passion. the job should not I believe consume their lives and being good at the job and a professional should not make the individual feel any less a good nurse by not standing up against others bad attitude or practice.	Service user 1	The comments from the service user do not necessarily represent the expectations of the professional body concerning promotion of professional image	None

	First statement is a bit lengthy and complex - needs breaking up	Academic 3	Appropriate comment	The statement has been broken up into two sentences and reads better.
	There are aspects i strongly agree with, but this is all dependant on how far they are in their training i.e. 1st year - agree, 3rd year strongly agree.	Darzi fellow 1	This is for third year which means the respondent strongly agrees.	None
	Think it is difficult to speak out maybe reports instead of challenges	Nurse Education Lead 1	Report is an expectation. Challenge is exceeding the expectations.	None
	The ability to challenge may be dependent on them having confidence and part of programme they are on.	Senior Nurse Education 1	That's why it was placed in exceeds expectations level.	None
6	I wonder whether you would like to add a mention of following chaperone policy	Practice educator 4	This implied in the statements	None
	Add in formal use of names and pet names?!	Academic 2	This competency statement is about dignity in relation to privacy. Being respectful and courteous is covered in Q7 above.	None
	Same comment as before	Commissioner 1	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	Re: Love to see: in an ideal world when time not an issue. Otherwise I feel all patients should have the same level of privacy and dignity and practice should take into account of this for all.	Service user 1	That's why it was placed in exceeds expectations level.	None
	Love to see experienced students leading junior students	Practice educator 1	Implied in stating "shapes their own care and that of others to ensure this is maintained"	None
	In the 'Expect to see' 'shows consideration' I would like to see 'and demonstrates in practice consistently'. I think 'Identifies challenges to privacy and dignity' would also be expected.	Practice educator 3	Demonstrates in practice consistently' is appropriate. The "exceeds expectation" level talks about being "thoughtful" about the challenges implying creativity in shaping practice.	The competency statement will read "demonstrates knowledge and shows consistency in maintaining dignity and privacy in their practice..."
	Whilst I agree with what is in the first statement it only addresses privacy - dignity is more than privacy it is about being treated as a valued individual. Second statement is woolly and not specific	Academic 3	Treating patients as a valued individual is part of Q7. This statement is specific to dignity in relation to privacy. The statement has been reworded as per the previous comment.	None
	As in previous - may be difficult to challenge	Nurse Education Lead 1	Challenge is exceeding the expectations.	None

7	Same comment as before	Commissioner 1	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	Expect to see. I would like to include 'open to constructive feedback and reflects on experiences to develop further'.	Practice educator 3	This is more appropriate in Q13	None
	Last statement: first part is woolly - what is professional conduct? Doesn't seem well related to the statement	Academic 3	The rest of the sentence explains what is meant by "professional conduct"	None
	Uncomfortable with the notion of not trusting students in the first part of the question, if we don't trust these people how can we build honesty into their learning	Darzi fellow 1	Appropriate comment	The statement now reads "Not perceived reliable to..." instead of "Not perceived trusted to..."
	If not able to act themselves than to be confident in their ability to report to their mentor and ward manager to action.	Nurse Education Lead 1	This is mentioned	None
8	Again dependant on year of student	Academic 1	The rubric is designed for 3 rd year students and this was made clear in the information provided.	None
	Same comment as before	Commissioner 1	The comment shows the respondent having difficulties in understanding how to respond to negative statements.	None
	Only agree with first statement as passive learning not necessarily negative. Doesn't always correlate with 'no learning'.	Other (joint academic/clinical commissioner) 1	The competency statement is related to making effort to engage in the learning, being passive means the student did not meet expectations. However, "passive learner" can be removed without affecting as the rest of the statement which provide a clear description.	Passive learner removed
	Feel passive learner is the wrong phrase to use	Darzi fellow 1	See above	See above
	This would be dependent upon student's own level of knowledge and aptitude in this area of nursing.	Academic 4	The rubric is designed for 3 rd year students and this was made clear in the information provided.	None
	Always use evidence based practice and challenge any poor practices observed.	Senior nurse education 1	This is included	None
	Terminology of the ranking. The majority (agree and strongly agree) favoured the meets expectations format (83.6%) over the others (love to see 53%, competent 51%, achieved 46.9, numerical 25%)		There were no comments on this section despite some of the participants are experienced educationalist and could have commented on the difference in terminology and provided more value.	The ranking "Does not meet expectations -Meets expectations - Exceeds expectations" will be used in the final rubric.

Appendix 18: Mentors questionnaire in phase two.



Stage 2- Mentor version

Page 1: Welcome to the questionnaire

Thank you for participating in the study. Completing the questionnaire should take under 10 minutes

If you have any questions please do not hesitate to contact me.

Ibraheim Almalkawi

Email: [redacted].ac.uk

Phone: 020 [redacted]

Page 2: Clarity and understanding of criteria and language

1. The rubric helped me recognise the required level the student needs to achieve.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

2. The rubric provided clear language to describe the levels.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

3. The rubric helped me explain to the student what is expected of him/her.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 3: Rigor of assessment

4. The rubric allowed me to distinguish between the different levels of competence.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

5. The rubric reduced my confusion around the required level to pass.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

6. The rubric provided adequate direction to me to understand and assess students' attitudes and behaviours.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

7. The rubric provided a clear guide to measure student performance accurately.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

8. The rubric gave me confidence that the decision and grading of the student is fair.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 4: Feedback

9. The rubric provided a structured approach to identify and plan the student's learning needs.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

10. The rubric helped me give accurate feedback that supported the level awarded.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 5: Mentor-student professional relationship and interaction

11. The rubric guided me to reflect on practice and real life situations with the student.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

12. The rubric was a useful tool to encourage reflective discussion between me and the student about his/her professional attitude and behaviour.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

13. Overall, I found the rubric practical to use.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 6: Reliability of the rubric

14. The use of the rubric will help me to be consistent in identifying and grading students' level of performance in the future.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

15. The rubric will ensure mentors involved in student assessment are consistent in accurately identifying the required level of competence.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 7: In the following questions, please feel free to share your thoughts about using the rubric.

16. What did you like best about using the rubric?

17. What did you like least about using the rubric?

18. Overall, did you find the rubric useful?

19. Any other comments you would like to add?

Page 8: Thank you very much for your contribution to this research.

Appendix 19: Students questionnaire in phase two.

Stage 2- Student version

Page 1: Welcome to the questionnaire

Thank you for participating in the study. Completing the questionnaire should take under 10 minutes

If you have any questions please do not hesitate to contact me.

Ibraheim Almalkawi

Email: [redacted].ac.uk

Phone: 020 [redacted]

Page 2: Clarity and understanding of criteria and language

1. The rubric helped me recognise the mentor’s expectations of the required level I need to achieve.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

2. The rubric provided clear language to describe the levels.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 3: Rigor of assessment

3. The rubric allowed me to distinguish between the different levels of competence.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

4. The rubric increases understanding around the required level to pass.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

5. The rubric provided adequate direction to help me understand the required level of attitudes and behaviours.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

6. The rubric allowed me to see the fairness of the decisions made to grade my level of performance.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 4: Feedback

7. The rubric provided a structured approach to identify and plan my learning needs.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

8. The rubric allowed me to self-assess my own level of performance.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 5: Mentor-student professional relationship and interaction

9. The rubric guided me to reflect on practice and real life situations with my mentor.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

10. The rubric was a useful tool to encourage reflective discussion between me and the mentor about my professional attitude and behaviour.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

11. Overall, I found the rubric practical to use.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 6: Reliability of the rubric

12. The rubric will ensure all mentors involved in assessing students will be consistent in accurately identifying the required level of competence.

	1	2	3	4	5	
Strongly agree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly disagree

Page 7: In the following questions, please feel free to share your thoughts about using the rubric.

16. What did you like best about using the rubric?

17. What did you like least about using the rubric?

18. Overall, did you find the rubric useful?

19. Any other comments you would like to add?

Page 8: Thank you very much for your contribution to this research.

Appendix 20: HEI ethical approval for phase two



Direct line: 020 [redacted]
E-mail: [redacted].ac.uk
Ref: hscsep/16/1

Mr Ibraheim Almalkawi

Friday 1st July 2016

Dear Ibraheim

RE: Testing a rubric for assessment of pre-registration nursing students in practice.

Thank you for submitting this proposal and for your amendments to the Participant Information Sheets, Consent form and Invitation E-mail made in response to the reviewers' comments.

I am pleased to inform you that full Approval for this study has been given by Dr [redacted] on behalf of the School of Health and Social Care School Ethics Panel.

If you wish to make any changes to the research protocol or any of the documents related to this study you MUST seek approval from this panel before making those changes.

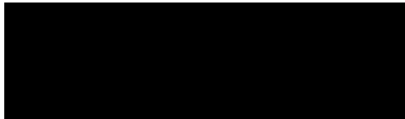
Please include your reference number (hscsep/16/1) in any future correspondence.

I wish you every success with your research.

Yours sincerely,

Dr [redacted]

Associate Professor
Chair, School Ethics Panel
School of Health and Social Care



Appendix 21: Permission from Head of Department at the HEI to access students

From: [REDACTED]
Sent: 17 September 2016 18:16
To: Almalkawi, Ibraheim
Subject: Re: Permission to access students

Ibraheim,

Very happy to support.

Regards

xxxxxx

On 17 Sep 2016, at 14:38, "Almalkawi, Ibraheim" [REDACTED] [.ac.uk](#)> wrote:

Dear [REDACTED]

I am a third year PhD student in the School of HSC undertaking a research project titled: **Testing a rubric for assessment of pre-registration nursing students in practice**. The aim of the research is to develop descriptors for the current statements in the practice assessment document, and then test whether or not the rubric improves the quality of assessment in practice placements.

The stage of developing the rubric in consultation with stakeholder through a Delphi process is now completed. The objective of the study at this stage is to test the rubric in practice.

As the Head of Department for Adult Nursing pre-registration programme at xxxxx, I would be grateful if you would permit me access to 3rd year students on the BSc (Hons) Adult Nursing to ask them to participate in the research. The plan is to recruit 30 students.



The School Ethics Panel in the School of Health and Social Care have reviewed and approved this study (Reference number HSCSEP/16/1), and am happy to send on the information sheet for the study or answer any questions if needed.

If you would be willing to permit access to students please email me back at [REDACTED] [.ac.uk](#) at your earliest convenience.

Yours sincerely,

Ibraheim Almalkawi
Part time PhD student
School of Health and Social Care

Appendix 22: NHS REC online tool with the decision that the study does not need ethical approval.

Do I need NHS REC approval?

i To print your result with title and IRAS Project ID please enter your details below:

Title of your research:

IRAS Project ID (if available):

Your answers to the following questions indicate that you do not need NHS REC approval for sites in England. However, you may need other approvals.

You have answered **YES** to: Is your study research?

You answered **NO** to all of these questions:

Question Set 1

- Is your study a clinical trial of an investigational medicinal product?
- Is your study one or more of the following: A non-CE marked medical device, or a device which has been modified or is being used outside of its CE mark intended purpose, and the study is conducted by or with the support of the manufacturer or another commercial company (including university spin-out company) to provide data for CE marking purposes?
- Does your study involve exposure to any ionising radiation?
- Does your study involve the processing of disclosable protected information on the Register of the Human Fertilisation and Embryology Authority by researchers, without consent?
- Is your study a clinical trial involving the participation of practising midwives?

Question Set 2

- Will your study involve research participants identified from, or because of their past or present use of services (adult and children's healthcare within the NHS and adult social care), for which the UK health departments are responsible (including services provided under contract with the private or voluntary sectors), including participants recruited through these services as healthy controls?
- Will your research involve collection of tissue or information from any users of these services (adult and children's healthcare within the NHS and adult social care)? This may include users who have died within the last 100 years.
- Will your research involve the use of previously collected tissue or information from which the research team could identify individual past or present users of these services (adult and children's healthcare within the NHS and adult social care), either directly from that tissue or information, or from its combination with other tissue or information likely to come into their possession?
- Will your research involve research participants identified because of their status as relatives or carers of past or present users of these services (adult and children's healthcare within the NHS and adult social care)?

Question Set 3

- Will your research involve the storage of relevant material from the living or deceased on premises in the UK, but not Scotland, without an appropriate licence from the Human Tissue Authority (HTA)? This includes storage of imported material.
- Will your research involve storage or use of relevant material from the living, collected on or after 1st September 2006, and the research is not within the terms of consent from the donors, and the research does not come under another NHS REC approval?
- Will your research involve the analysis of DNA from bodily material, collected on or after 1st September 2006, and this analysis is not within the terms of consent for research from the donor?

Question Set 4

- Will your research involve at any stage intrusive procedures with adults who lack capacity to consent for themselves, including participants retained in study following the loss of capacity?
- Is your research health-related and involving prisoners?
- Does your research involve xenotransplantation?
- Is your research a social care project funded by the Department of Health?

Appendix 23: Approvals from individual Trusts to access mentors

Trust 1

From: [REDACTED] Hospitals [REDACTED].nhs.uk]
Sent: 18 September 2016 16:33
To: Almalkawi, Ibraheim
Cc: [REDACTED] Hospitals
Subject: RE: Permission to access mentors

Dear Ibraheim,

Thank you for coming to see me yesterday to talk through this study. I would be happy to support your research at [REDACTED] and for you have access to the mentors of students for the purposes of evaluation of your thesis.

Good luck with it, I look forward to hearing the outcome.

I have cc'd [REDACTED] to this email as one of the Registered Workforce Team so she can make that team aware as they work with the mentors and students.

Kind regards

[REDACTED]
[REDACTED] | **Associate Director of Education & Training**
Education, Training, Learning & Development | Corporate Division
[REDACTED] NHS Trust

I:
E:

From: Almalkawi, Ibraheim [mailto:[REDACTED]@ac.uk]
Sent: 17 September 2016 14:22
To: [REDACTED]
Subject: Permission to access mentors

Dear [REDACTED]

I am a third year PhD student in the School of Health and Social Care at xxxxxx University undertaking a research project titled: **Testing a rubric for assessment of pre-registration nursing students in practice**. The aim of the research is to develop descriptors for the current statements in the practice assessment document, and then test whether or not the rubric improves the quality of assessment in practice placements.

The stage of developing the rubric in consultation with stakeholder through a Delphi process is now completed. The objective of the study at this stage is to test the rubric in practice. This will involve both students and their mentors using the rubric throughout the placement then be asked to complete questionnaire.

The plan is to recruit 30 third year students from the Adult Nursing field and pair them with their allocated mentors. Once 30 are recruited, their named mentor will be invited to take part too. I would be grateful if you would permit me to invite mentors through Trust emails.

It is unclear at this stage how many mentors will be recruited from your Trust as it will depend on the number of students placed in your Trust who agreed to participate. The School Ethics Panel in the School of Health and Social Care at xxxxxx University have reviewed and approved this study (Reference number HSCSEP/16/1), and am happy to send on the information sheet for the study or answer any questions if needed.

If you would be willing to permit access to mentors please email me back at [REDACTED].ac.uk at your earliest convenience.

Yours sincerely,

Ibraheim Almalkawi
Part time PhD student
School of Health and Social Care
xxxxx University

From: [REDACTED].nhs.uk]

Sent: 15 October 2016 12:45

To: Almalkawi, Ibraheim

Cc: [REDACTED]

(RGC) Theatre Practice Facilitator

Subject: RE: Permission to access mentors

Thanks ibraheim

It was good to meet you too

As discussed on Thursday probably the easiest way to coordinate this will be directly through [REDACTED] ccd above.

We will identify the mentors in the next week - dawn will be in touch to agree the next steps

Hope that is ok

Please let us know if there is anything else we should be doing

BW [REDACTED]

[REDACTED]
Deputy Director
Education Academy

From: [REDACTED] nhs.uk [REDACTED] nhs.uk]
Sent: 27 October 2016 16:08
To: Almalkawi, Ibraheim
Subject: FW: Permission for a project

Dear Ibrahiem

I have read through your documents and believe this is as you state a service evaluation project not requiring NHS Ethics approval. As this is a service evaluation rather than research you will need approval in this case as you say the education lead at [REDACTED]. As the study is not research we do not need to review the documents.

I am content that our review is not required and would be content for the education Lead of [REDACTED] to provide their approval

Best wishes

[REDACTED]
Research Portfolio Coordinator
Joint Research Office
(part of the Research Support Centre)

From: [REDACTED] .nhs.uk [REDACTED] .nhs.uk]
Sent: 28 October 2016 13:37
To: Almalkawi, Ibraheim
Cc: [REDACTED] nhs.uk
Subject: RE: Permission for a project

Dear Ibraheim

As stated in the email from [REDACTED] approval has been given by [REDACTED] for you to conduct your research. As the Lead nurse for pre-registration at [REDACTED] I can confirm that we are happy to support your research and approach mentors. We hope that your research will inform on how we continue to support mentors and students in practice.

I have CC'd in [REDACTED] who is our Deputy Chief Nurse for pre and post education and was informed of the original request.

Best wishes

[REDACTED]

Lead nurse pre-registration education (RN, Bsc, PGCE, MA Practice Education)

From: [REDACTED] .nhs.uk]
Sent: 14 November 2016 15:49

To: Almalkawi, Ibraheim
Subject: RE: Permission to access mentors

Dear Ibraheim,

Thank you for your email and apologies for my delay in responding. I am happy for you to approach mentors at [REDACTED] NHS Trust to be able to carry out your research. It appears from your email that you do not need to apply Trust ethical approval to be able to do this. However, if you need further assistance please don't hesitate to get in touch.

Yours

[REDACTED]

ET&D Manager – Clinical Education

Education, Training & Development

From: [REDACTED] [ac.uk](#)>
To: "Almalkawi, Ibraheim" <almalkai@xxxx.ac.uk>
Cc: [REDACTED]
Subject: Research at xxxx
Date: Fri, Nov 11, 2016 18:18

Hi Ibraheim,

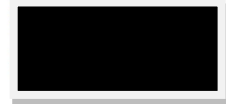
Hope you are well. I visited the students at [REDACTED] today to discuss their participation in the your PhD research. All say they are still willing. I spoke with [REDACTED] (Lecturer in Palliative Care [REDACTED] who liaises with the students) and discussed your research with her. I have also copied [REDACTED] into this as [REDACTED] is development nurse for the hospice. I provided copies of the information sheet, consent forms and rubric for the mentors and [REDACTED] has kindly agreed to meet with the mentors early next week to disseminate the information and gain consent. Please contact [REDACTED] if you have any queries about this.

Hope that is ok for you and let me know if you require any additional help

Appendix 24: Indemnity letter from the HEI.



School of Health and Social Care






17th October 2016


Dear Sir

Title of study: Testing a rubric for assessment of pre-registration nursing students in practice

Name of applicant: *Ibraheim Almalkawi*

 UNIVERSITY is willing to take on the role of sponsor in relation to this research project, to be carried out by Ibraheim Almalkawi who is currently a student studying for PhD. The research study is part of that award. The academic supervisors for the project are Professor  and Dr .

I confirm that indemnity will be in place covering this project the details of the cover are in the attached statement of insurance cover.

All correspondence for the sponsor should be marked for the attention of Professor .

Yours sincerely,



Professor of Clinical Nursing Practice



Appendix 25: Participants information sheet for phase two (students)

Reference number: HSCSEP/16/1

Participant Information Sheet- students

Title: Testing a rubric for assessment of pre-registration nursing students in practice

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish. Please contact me if you need further clarification or would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of this study?

The purpose of this study is to test a scoring rubric that is designed to interpret the professional values competency statements in the PAD currently in use. The aim is to evaluate whether the rubric with clearly described levels of achievements improves mentors' effectiveness and confidence in judgement decisions as well as enhancing learning by providing clearly structured feedback.

Why have I been invited to take part?

The rubric is designed to interpret the professional values for third year students in the field of adult nursing. You are being invited to participate as you are a 3rd year adult nursing student and it is important to elicit your views on the effectiveness of the rubric. I am aiming to recruit 30 students and their mentors. Participation in the study necessitates both the students and their mentors are willing to take part.

Taking part in the study is voluntary, and you are free to withdraw at any time without the need for explanation and you will not be disadvantaged.

What will I be asked to do if I take part?

If you are willing to participate, I will meet with you and your mentor together in the placement area to explain how the rubric should be used and ask you to sign an informed consent form. Then both you and your mentor will be asked to use the rubric alongside the PAD throughout the placement. The rubric is designed to interpret levels for half the professional values statements only. The remaining statements in the PAD will be completed using the usual method. This will provide you with the opportunity to experience the assessment process with and without using the rubric. When the placement is finished, you will be asked to complete an online questionnaire which should take approximately 10 minutes to complete. Your response would be required within one week of the end of the placement. A reminder email will be sent to participants a week later.

What are the possible benefits and risks of this study?

It is not anticipated that you will be disadvantaged or suffer any risk from this study. The rubric is to be used as an adjunct to the PAD rather than to replace it; therefore, the assessment will follow the university procedures as specified in the PAD and your marks on placement will not be influenced by your participation in the study. However, if you experience any distress please remember that you are free to withdraw from the study at any time without the need for explanation, and can contact me for support.

It is unlikely that you will gain any personal benefit from participating in this research, but your input is valuable in attempting to improve the accuracy of assessment of student nurses in the practice setting.

Will my talking part in this study be kept confidential?

Data will be collected through a secured website where only invited members will be allowed access. All information received from you will be collated anonymously, handled in a confidential manner and stored in a locked filing cabinet and on a password protected computer in an environment locked when not occupied. All responses received in the study will be strictly confidential, and your identity will not be disclosed. Direct quotes to free-text answers may be used in the study report, but will not be traceable back to you. Should any issue of concern be revealed in the questionnaire, I will discuss the issue with you and together decide whether to raise the concern.

Only me and my supervisors will have direct access to the information. Any reference to you will be coded. This information will be held for a maximum of 5 years after the PhD project is completed and then it will be destroyed. You will have the right to access submitted information according to the UK data protection laws.

Who has reviewed the study?

This study is being completed as part of a PhD degree at xxxxx University. It has been reviewed and ethically approved by the School Ethics Panel in the School of Health and Social Care.

Who do I contact for more information or if I have concerns?

If you wish for any further information or you have concerns about any aspect of this study or experience any distress you can stop completing the survey and feel free to speak me. If you have any complaints about the way you have been dealt with during the study or other concerns you can contact Dr. [REDACTED] at [REDACTED] who is the Academic Supervisor for this study. Finally, if you remain unhappy and wish to complain formally, you can contact the Chair of the School of Health and Social Care Ethics Panel. Details can be obtained from the university website: [https:// \[REDACTED\].ac.uk/page/research-degrees-ethics](https://[REDACTED].ac.uk/page/research-degrees-ethics)

What do I do now?

Thank you for reading this information sheet and for considering taking part in this research. Please respond to this email within a week to let me know if you wish to take part in the study.

If you have any questions or concerns please do not hesitate to contact me.

Ibraheim Almalkawi
PhD Student

[REDACTED]
[REDACTED]
[\[REDACTED\].ac.uk](#)

Professor [REDACTED]
Director of studies

[REDACTED]
[\[REDACTED\].ac.uk](#)

Dr. [REDACTED]
Supervisor

[REDACTED]
[REDACTED]
[\[REDACTED\].ac.uk](#)

Appendix 26: Participants information sheet for phase two (mentors)

Reference number: HSCSEP/16/1

Participant Information Sheet- Mentors

Title: Testing a rubric for assessment of pre-registration nursing students in practice

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about this study if you wish. Please contact me if you need further clarification or would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of this study?

The purpose of this study is to test a scoring rubric that is designed to interpret the professional values competency statements in the PAD currently in use. The aim is to evaluate whether the rubric with clearly described levels of achievements improves mentors' effectiveness and confidence in judgement decisions as well as enhancing learning by providing clearly structured feedback.

Why have I been invited to take part?

The rubric is designed to interpret the professional values for third year students in the field of adult nursing. You are being invited to participate as you are the allocated mentor to a 3rd year adult nursing student who agreed to take part in this study. I am aiming to recruit 30 students and their mentors, and participation in the study necessitates both the students and their mentors are willing to take part.

Taking part is voluntary, and you are free to withdraw from the study at any time without the need for explanation and neither you nor the student will be disadvantaged.

What will I be asked to do if I take part?

If you are willing to participate, I will meet with you and your student together in the placement area to explain how the rubric should be used and ask you to sign an informed consent form. Then both you and the student will be asked to use the rubric alongside the PAD throughout the placement. The rubric is designed to interpret levels for half the professional values statements only. The remaining statements in the PAD will be completed using the usual method. This will provide you with the opportunity to experience the assessment process with and without using the rubric. When the placement is finished, you will be asked to complete an online questionnaire which should take approximately 10 minutes to complete. Your response would be required within one week of the end of the placement. A reminder email will be sent to participants a week later.

What are the possible benefits and risks of this study?

It is not anticipated that you will be disadvantaged or suffer any risk from this study. The rubric is to be used as an adjunct to the PAD rather than to replace it; therefore, the assessment will follow the university procedures as specified in the PAD and students marks on placement will not be influenced by their participation in the study. However, if you experience any distress please remember that you are free to withdraw from the study at any time without the need for explanation, and can contact me for support.

It is unlikely that you will gain any personal benefit from participating in this research, but your input is valuable in attempting to improve the accuracy of assessment of student nurses in the practice setting.

Will my talking part in this study be kept confidential?

Data will be collected through a secured website where only invited members will be allowed access. All information received from you will be collated anonymously, handled in a confidential manner and stored in a locked filing cabinet and on a password protected computer in an environment locked when not occupied. All responses received in the study will be strictly confidential, and your identity will not be disclosed. Direct quotes to free-text answers may be used in the study report, but will not be traceable back to you. Should any issue of concern be revealed in the questionnaire, I will discuss the issue with you and together decide whether to raise the concern.

Only me and my supervisors will have direct access to the information. Any reference to you will be coded. This information will be held for a maximum of 5 years after the PhD project is completed and then it will be destroyed. You will have the right to access submitted information according to the UK data protection laws

Who has reviewed this study?

This study is being completed as part of a PhD degree at xxxxx University. It has been reviewed and ethically approved by the School Ethics Panel in the School of Health and Social Care.

Who do I contact for more information or if I have concerns?

If you wish for any further information or you have concerns about any aspect of this study or experience any distress you can stop completing the survey and feel free to speak me. If you have any complaints about the way you have been dealt with during the study or other concerns you can contact Dr. [REDACTED] at [REDACTED] who is the Academic Supervisor for this study. Finally, if you remain unhappy and wish to complain formally, you can contact the Chair of the School of Health and Social Care Ethics Panel. Details can be obtained from the university website: [https://\[REDACTED\].ac.uk/page/research-degrees-ethics](https://[REDACTED].ac.uk/page/research-degrees-ethics)

What do I do now?

Thank you for reading this information sheet and for considering taking part in this research. Please respond to this email within a week to let me know if you wish to take part in the study.

If you have any questions or concerns please do not hesitate to contact me.

Ibraheim Almalkawi
PhD Student

[REDACTED]
[REDACTED]
[\[REDACTED\].ac.uk](mailto:[REDACTED].ac.uk)

Professor [REDACTED]
Director of studies

[REDACTED]
[\[REDACTED\].ac.uk](mailto:[REDACTED].ac.uk)

Dr. [REDACTED]
Supervisor

[REDACTED]
[REDACTED]
[\[REDACTED\].ac.uk](mailto:[REDACTED].ac.uk)

Appendix 27: Informed consent form for participating in phase two

Reference number: HSCSEP/16/1

Title: Testing a rubric for assessment of pre-registration nursing students in practice

	Tick
I have read the attached information sheet about the research in which I have been asked to participate and have been given a copy to keep. I have had the opportunity to discuss the details and ask questions about this information.	
The Investigator has explained the nature and purpose of the research and I understand what is being proposed.	
I understand that my personal involvement and my particular data from this study will remain strictly confidential.	
I understand that my identity will be protected and any quotes used will be coded.	
I have been informed about what the data collected in this investigation will be used for and how long it will be retained.	
I understand that I am free to withdraw from the study at any time without given any reason until data is anonymised and irretrievably merged with others.	

I hereby fully and freely consent to participate in the study.

Participant's name:

Participant's signature:

Date:

As the main researcher responsible for this study I confirm that I have explained to the participant named above the nature and purpose of the research to be undertaken.

Name:

Researcher's signature:

Date:

Appendix 28: A reminder email to participants in phase two after one week

Dear participant

I recently invited you to complete a short questionnaire about your experience using the rubric. If you have already completed the questionnaire, thank you very much and please accept my apology for sending this reminder.

Your response is very important to establish the usefulness of the rubric. Hence, I am writing to encourage you to complete the questionnaire.

[Click here to begin the questionnaire](#)

Kind regards

Ibraheim Almalkawi (PhD Student)

School of [REDACTED] | [REDACTED] University

t: +44 (0)20 [REDACTED] e: [REDACTED]@xxxx.ac.uk

Appendix 29: A final email reminder to participants in phase two after two weeks

Dear participant

If you have already completed the questionnaire, ignore this email and thank you very much and please accept my apology for sending this reminder.

This is just a final prompt for those still wanting to do it that it's not too late!

This questionnaire will take around 5 minutes to complete.

Your participation in this study is very valuable, I appreciate you taking time from your busy schedule to contribute to this study.

[Click here to begin the questionnaire](#)

Kind regards

Ibraheim Almalkawi (PhD Student)

School of [REDACTED] | [REDACTED] University

t: +44 (0)20 [REDACTED] e: [REDACTED]@xxxx.ac.uk

Appendix 30: Key messages from the qualitative responses in phase two coded according to their category

Categories	Mentors feedback	Students feedback
Clarity of the language	<ul style="list-style-type: none"> • More easy to use in assessment of students. • Its clarity (liked most). • Easy to understand. • It was easy to use, and much clarity in reducing any conflicts or misunderstanding. • Gave a structure. • It gives clear and concise view and plan. • Very simple and straight forward to use with clear instruction. • Provided clarity & structure. • Simple to use and understand. • It's well structured and easy to use. • It makes it clearer in understanding the student's packs. • Easy to understand and made scoring of student more straight forward and was easy to explain to student. • The tool set out a clear guideline. • The words to put up while writing about the student objective. • Very informative in helping to choose the right word in completing students pack. • Helps to use the right language. • Lots of information (like most). 	<ul style="list-style-type: none"> • Easy to follow. • Easier for my mentor to assess me. • Simplicity in details. • It was informative and easy to use. • It was easy to follow and understand. • It's simplicity, largely uncomplicated and easy to follow. • Having a clear and fair guideline. • Clear and comprehensible. • In the past the wording of the PAD itself and the skills to achieve was daunting. • Easy to understand. • Clear and simple to understand. • The clear language used to express expectations. • Clear and simple language. • Clearly structured and worded. • The language was very clear. • Sometimes the concepts in the PAD can be abstract and vague and it helped give more clarification. • It gives clarity of what is expected of you. Hence helps both the student and mentors

Distinction between levels of competence	<ul style="list-style-type: none"> • Clear grading system for the student. • Provided a clear guideline of what was expected of the student. • Easy to follow in grading student. • This gives me an idea to assess the student and grade the student. • Found it very useful as a guidance to distinguish between 'fail/pass/and exceptional performance'. • It gave me a clearer insight into what was expected of the student and how to grade them. • I also liked that it allowed me to say whether they were achieving expectation or merely meeting it - much better than just saying achieved or not achieved. 	<ul style="list-style-type: none"> • Practical examples to differentiate between competence levels. • Explained expected level. • The rubric made it easier to understand different levels of outcomes. • Guided my mentor throughout regarding expectations. • Enabled myself and mentor to understand the expectations needed. • It helped give a more extensive explanation of what was expected of me. • Sets out what exactly I need to do to achieve my assessments. • It helped me a lot to understand the required behaviour and attitudes. • It was very precise about what was expected of me unlike before where the values seemed vague. • It made understanding what is required of my competences simple. • Makes clear what is expected during placement. • It gives a clarity of what is expected of you • Guiding the mentor and as a student provided guidance as to the standards we should be reaching. • It gives clarity of what is expected of you. Hence helps both the student and mentors. • Gives mentors a basis for marking. • I think that the idea of fail, pass or exceeds is a great way to determine the difference between students who pass or excel during placement
Rigour of the decision-making	<ul style="list-style-type: none"> • ...so one do not assess a student on assumption. It also assists the facilitator in making their decisions. • Supports consistency, and evidencing assessment • It helps to assess the student. • It's clear and decision making is easier. • I think some facilitators can be very rigid in their opinion about students; this will give a uniform plan to assess students. 	<ul style="list-style-type: none"> • Having a clear and fair guideline. • Fairer for students. • I liked most about the rubric style of assessment is the fairness. It eliminates any worries that I will be judged on anything but my performance. • ...and gives me confidence that I will be judged entirely fairly. • It allowed fair and fast assessment. • Made the decision process smooth and more accurate. • This process allows mentors to assess students solely on their nursing skills. • It was a clear tool for the mentor that dispelled any issues of bias. • ...and more thought put into the marking.

	<ul style="list-style-type: none"> • Helping the students understanding and your own understanding of the level needed for passing the placements. (More standardised). • Assess students fairly and consistently. 	<ul style="list-style-type: none"> • The best thing is that it gets rid of any bias or unachievable expectations by a mentor. • Made the midpoints and finals more structured and more thought put into the marking.
Facilitation of learning, feedback and self-assessment	<ul style="list-style-type: none"> • The rubric allowed me to discuss in detail what was required of the student to achieve higher levels in each area. • ...It also allowed you to build an action plan to achieve excellence and not just when students are failing to meet standards expected. • I was able to communicate and reflect with the student on her entire placement. 	<ul style="list-style-type: none"> • It provided me with a target to exceed placement instead of just passing. • It gave me a sense of direction. • Also allows for professional development and recognises behaviours which are better than simply pass. • It provided guidelines for assessment and a target to achieve the exceptional status. • How I could improve in practice. • ...also helped me to assess my own progress. • It provided me with a standard of practice that I could recognise and build upon if necessary. • It gave the mentor a guide to the level expected and helped me to understand what she thought my strengths and weaknesses are. • Allowing me to use the rubric to reflect upon situations I have encountered and dealt with in my practice. • Allowed myself to comprehend the required elements to achieve each component of my professional values. • It helped me a lot to understand the require behaviour and attitude. • Explained the professional values thoroughly making it easier. • Both can agree on any other learning opportunities that may arise whilst in placement. • Also helps to have something to work towards. • It serves as a guide. • I found it useful on a personal level. •it makes it much easier to see what level you are at and how to improve or maintain your standards.

		<ul style="list-style-type: none"> • Really defines for students HOW they can achieve their professional values. Much better than the current statements which only tell you WHAT they are expecting. Allows us to work smarter rather than harder. • ...allowed for a deeper discussion into what exactly the professional value is. • Useful tool for future students and mentors as it provides a framework to guide to discussing and achieving professional values.
Time	<ul style="list-style-type: none"> • Save time. • Saves time in filling out paper work. • Tt helped with the flow of the students assessment and the interview was shorter which helps when you are busy. 	<ul style="list-style-type: none"> • Saves time for mentors. • It allowed fair and fast assessment. • Slightly more time consuming, but allowed for a deeper discussion into what exactly the professional value is. • It's more time consuming especially when you're on a busy unit. • It did make the process more time consuming but I don't know how that could be avoided. • Even though it's time consuming, it makes it much easier to see what level you are at and how to improve or maintain your standards.
Student-mentor interaction and dialogue	<ul style="list-style-type: none"> • Prompted good discussion between myself and the student. • I was able to communicate and reflect with the student on her entire placement. 	<ul style="list-style-type: none"> • Was helpful in discussion with mentors. • Both me and my mentors found it very useful as so many time as so many times in past the wording of the PAD itself and the skills to achieve was daunting. • My mentor and I sat and discussed each section together. • It allowed me and my mentor maintain continuity through my assessment. • This time round it wasn't a mentor giving you their version of what they think that means but working together with you to achieve what both of us know should be achieved. Awesome really. • It gave the mentor a guide to the level expected and helped me to understand what she thought my strengths and weaknesses are. • Both can agree on any other learning opportunities that may arise whilst in placement. • Help mentors and students on placement • Prevents conflicts and disagreements between mentors and students. • Provides a framework to guide to discussing and achieving professional values.

		<ul style="list-style-type: none"> • Very useful for some students who may experience personality clashes. • I can imagine a situation in which the mentor might doubt someone is achieving but the student would have evidence to demonstrate that they have achieved the expected requirements.
Participants reflection on using the scoring rubric	<ul style="list-style-type: none"> • Still needs total perfection. • Too much detail. • In some areas the examples were still somewhat vague and thinking of examples that could demonstrate excellence in practice was difficult. • Very useful I hope it comes into use and is used for all students • Not a finished complete tool yet, therefore hard to understand if the whole tool will be as effective. Would love to see the finished result. • Would be more helpful on the other values as they are more "wishy-washy". • It only covered a limited area. • Hope it will be added in all of the students packs. • I hope it comes in to use and is used for all students. • I hope this will be used as regular guidelines in mentoring students. 	<ul style="list-style-type: none"> • May not give a room to express your own findings and reasons for grading • I felt nervous as if every part of me was going to be scrutinised more because of how specific it was. • I felt I would be achieving a higher standard i.e. exceptional whereas my mentor might have not felt that I was at this level. • There were several times when the mentor said things like "but there are many nurses who don't do that". It seemed like many of the goalposts in the 'exceeds expectations' category were higher than what even the average nurse performs at. • I am glad to have used the rubric. • I think it would be fantastic for everyone to use....and I will recommend it. • Would be useful if implemented into practice for students in the future. • It was only for the first eight questions. Would have preferred the rubric for all the questions. • Would like rubric for all the competencies in future. • It should have been used for the entire section and not for just the eight questions. • It was very helpful but would be better if available for all criteria • Please make one for all criteria. • What about the skills? • Great use to all student nurses from 1st-3rd year!

