Workplace-based assessment in Chinese medicine: a case study viewing Competency-Based Medical Education through a Practice Theory Lens

Felicity Clare Moir

University College London, Institute of Education

Department of Culture, Communication and Media

Thesis submitted for the degree Doctor of Education

Nov 16th 2020

Declaration of Own Work

I, Felicity Moir, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Declaration of Word Count

The exact number of words in this thesis is 45,688. The abstract, statements, acknowledgements, glossary, list of tables and appendices, table of contents, references and appendices are excluded from the word count.

FMoir

Abstract

This thesis explores how workplace-based assessment (WBA) was being
conducted in three institutions in different Western countries training students to
be practitioners of Chinese Medicine. While competency-based medical
education (CBME) as a curriculum model is being globalised, the literature on
its application to WBA shows a tension between standardization and
authenticity and between atomization and holism. The aim of this thesis was to
understand how the constructs of CBME were being enacted and how they
might be constraining or enabling WBA in these institutions.

Using a multiple embedded case-study design I explored WBA through data
gathered from the relevant accreditation documents in each country,
institutional documents and interviews with managers and clinical supervisors.
My analysis revealed that the different stakeholders manifested quite different
stories of assessment. In the accreditation and institutional documents the
discourse of measurement predominated; in contrast, while the managers were
leaning towards standardization and objectivity, they were also aware of a more
complex assessment culture. For the clinical supervisors the psychometric
grades being enacted could not be seen as a legitimate measure of objectivity
as the authentic and holistic constructs of CBME dominated.

The thesis draws on Schatzki's 'practice turn' as an interpretive lens. By viewing
WBA as a socio-cultural-material, embodied and interactive phenomenon I
show that it was the inter-relationships between institutionalized and disciplinary

discourses; between standardized and personalized competencies; between educational and practitioner identities; and the entanglement of artefacts and spatio-temporal arrangements, that enabled or constrained how WBA was enacted.

Through this analysis I argue that an understanding of how different assemblances lead to different enactments opens up possibilities of how to develop legitimate and workable assessment processes not just for Chinese medicine but any health profession.

Impact Statement

Existing research in the field of competency-based medical education (CBME) highlights the tensions that are created when the standardization and atomization of competencies are melded with the authenticity and holism that occurs in workplace-based assessment (WBA) (Eva and Hodges, 2012; Govaerts, van der Vleuten and Holmboe, 2019). This study has questioned this dialectic and found that CBME is not value-neutral, and even though the socio-constructivist perspective is recognized it still remains rooted in an instrumental and technical epistemology that strives for impartiality and a pre-determined end. By investigating how institutions teaching Chinese medicine (CM) are dealing with CBME in their assessment of students' competence to practice three areas have been illuminated that make an original contribution to the literature.

Firstly, this study has expanded a very limited research in CM education in the West. A clinical assessment is constructed not only by the knowledge, capabilities, and values internal to its own traditions but also by practices external to those traditions, such as the meta-practices of CBME. What has been demonstrated is that it is the very discourse of CBME that needs consideration in order to harmonise assessment across stakeholders and support accountability. If an institution wants to introduce a new assessment instrument, the complexity and entwinement of all the factors that form the practice of assessment – cultural, material and social - need to be considered. I have already presented preliminary findings at a British Acupuncture Accreditation Board workshop.

> Felicity's work on clinical education is timely and addresses practical wranglings we're all having. Great to disseminate this to clinical teachers in such an immediate and interactive way (Feedback BAAB education day 28-06-2018).

The understanding I have developed over this thesis is supporting my role as an external examiner for courses in the UK and abroad and will inform a review of the Standards for Education and Training for Acupuncture of the BAcC in which I will be the educational lead.

Secondly, by introducing a new theorization that foregrounds assessment as a 'practice', the study has generated empirical evidence to show how the constructs of CBME are being enacted. What was unique in this study was that I did not just investigate the assessment instruments, but contrasted the meaning assessment held for accreditors, institutions, managers and supervisors. By so doing I show how it is the relationships between all the stakeholders, the artefacts, the environment in which they work and the discourses employed, that form the practice of WBA. This suggests why strategies to make assessment instruments and assessors more objective may not be working.

Thirdly, there is very little empirical evidence in the literature on practice theory and its application to the world of medical education. van Lankveld *et al.* (2017) investigated the role of teacher communities and teaching courses in strengthening teachers' identities, agency and inter-connectedness. Meddings (2017) used a phenomenological methodology to analyse marking and grading by health academics. This study supports and expands the use of practice theory to better understand the phenomenon of WBA.

Reflective Statement

Hamilton (2005) presents the following insight into the research process:

> In practice, I suggest, research is always a fumbling act of discovery, where researchers only know what they are doing when they have done it; and only know what they are looking for after they have found it (p. 288)

My experience as a researcher was naive and fumbling but in my act of fumbling there was discovery. It is the doing of research that is key to understanding it. This reflective statement will consider my process of becoming a researcher from the factors that led me to the EdD and the themes that provided coherence throughout. The aim in this reflection is to add to the suggestions on what enables the EdD journey and supports the development of our multiple professional roles.

My professional life has been focused on the practice of Chinese Medicine (CM) and the education and training of students to become practitioners of CM. In moving from founding and running a private school into a University the possibility and expectation of developing an academic career was made possible. I completed an MSc in Inter-professional Practice and found it a seminal experience intellectually and in support of my role as a clinical teacher, publishing two papers as a result of the research I conducted (Moir, 2003; Moir, 2005). What should have been a logical next move into doctoral level research was not realised, however, for another 9 years with two key issues triggering my interest. Firstly the need to have a doctorate was becoming more critical in terms of status both as an academic in an HE institution and as a practitioner of a discipline that was still marginalised. Secondly, with my work more focused on management, I needed a different intellectual challenge and the structure of the doctorate would provide, and has provided, that. The decision to focus on education in Chinese medicine came from my personal passion and that I felt my 40 years of experience gave me something to offer. The topic of MOE1, 2 and the IFS of the EdD was related to a master's level module I was running at the University about the clinical reasoning process of CM. The focus of the thesis turned to the use of competencies in the clinical assessment of our

students. The place of competencies within workplace-based assessment (WBA) has been a key part of my role and I was one of a team of practitioner-educators who developed the Standards of Education and Training for Acupuncture (BAcC, 2011) for my profession. It was the EdD, however, that opened my eyes to how the cultural-discursive, material-economic and social-political arrangements that emanated from the policies and procedures of contemporary education were impacting on assessment frameworks. It was the literature I came across in the EdD that allowed me to see WBA 'strangely' (Kuper, Whitehead and Hodges, 2013).

The following three sections describe my journey through the EdD as a coherent process, namely pattern recognition, criticality and embodiment.

Pattern recognition

Chinese medicine is about recognising patterns in the signs and symptoms presented by patients. How a practitioner both perceives and understands those patterns is determined by their propositional, professional and personal knowledge, a dialectical process that blends the knowledge from the canons of the discipline and also their experience of practice. Research, I argue, is similar. In applying this theme to the research process, I can see that when I embarked on the EdD I neither had the breadth of propositional knowledge about my topics, about research methodology or conceptual frameworks, nor the professional experience to draw on in order to recognize patterns. Thus the early years of the EdD were spent learning a quite new sociological literature with which to understand education and a range of research methodologies. The taught modules introduced me to a wide variety of writers and ideas, the following standing out as key. Lorna Unwin's lectures on how approaches to pedagogy and assessment have changed over time and Norman Lucas's critique of how decontextualized standards are applied across diverse contexts were, ultimately, instrumental in the change of direction from the IFS to my thesis. The work of Stephen Ball and his critique of the neo-liberal agenda now pervading HE helped me to look differently at competency-based education. In my role as a 'middle leader' (Grootenboer, Ronnerman and Edwards-Groves,

2017) I came to question the ubiquitous use of competencies to direct a curriculum and the strait jacketing that was being expected with the mapping of learning outcomes onto assessments. The lecture by Paul Temple and an introduction to Edwards and Usher's approach to space and the curriculum sowed seeds that I only now realize emerged in my thesis. The Philosophy of Education seminars I attended introduced me to Michael Young's critique of powerful knowledge, Wittgenstein on how meaning is negotiated and Vygotsky's language as tool and were all key to shifts in my perspective and can be seen as threads through my thesis. The Medical Education seminars opened me to the ethnographic work of Paul Atkinson, Bruno Latour and Annemarie Mol and new approaches to conducting rigorous qualitative research. An immersion in this diverse literature provided the propositional knowledge I needed and through which I was able to start seeing the beliefs and opinions about education and research that I was taking-for-granted. Acknowledging one's biases are key to any pattern recognition. It is not that one can step outside them but can foreground them so as to check assumptions and conclusions being drawn.

Criticality

The iterative structure of the EdD was important in providing the opportunity to apply the whole research process, from conceptual design through to implementation, on a small-scale project before embarking on the larger thesis. As part of the process, feedback on the various assignments from the taught courses and IFS were what supported my development as a researcher.

A recurring comment on my assignments was the lack of application of a critical voice to my arguments. My interest was in my findings and the implications on practice and teaching rather than the development of the skills needed to be a critical researcher. On reflection, my problem was coming from two angles. One was due to my not understanding the literature because of a lack of wide enough reading and so was lacking the iterative relationship between the literature and the application. I was challenged on my ideological and conceptual positions, and as said before, the biases that go with that. Secondly,

I did not have my own research questions clearly thought through. I realised, however, that I had a problem with the latter. My interpretation was that having such a tight framing for my questions assumed a deductive approach to the literature while maintaining a broad view allowed for the inductive to emerge. While a tighter framing can limit the problem space, my concern was that it could also lead to premature closing. For example in the feedback for my thesis proposal the (too) many possible interpretations that the readers thought I was proposing for my aims were all, as far as I was concerned, valid and areas I wanted to pursue. However, finding a way to limit the area of interest so that I could go into more depth was key to focusing my research and that what I was lacking was a critical reading of the literature that could help illuminate what was important.

Embodied knowing

Hopwood (2016, p. 54) proposes that 'the value of theory is only ever tangible when it becomes entangled with data, with the empirical'. Hopwood's insight, was made concrete in my experience of the thesis. While I had read about Schatzki's 'practice turn' I had not understood it as a theory to apply to my data until I had gathered the data and was attempting to analyze it and understand it. The thesis started as an investigation into CBME and the tensions between the instrumental-technical and socio-constructive perspectives on WBA. Practice theory emerged from and formed the framework of the thesis only as I became immersed in the data gathering and analysis and was attempting to understand my findings.

A recommendation from my IFS was crucial to my progression in the thesis.

> *Jumping to analytic codes without descriptive codes could lead to errors in analysis and will make process of analysis less transparent (MN)*

As I found I was forming analytic codes while in the moment of interviewing, it was repeated and extensive descriptive coding that was needed to mitigate a premature interpretation. I also took Braun and Clarke's (2006) recommendation to write rich individual case reports for each interviewee. In

this way the coding remained linked to the individual story maintaining a relational quality. To gain 'analytic bite' Atkinson (2017, p. 2) proposes that clear and sustained analysis is needed with the movement between seeing the detail and the whole an important step in this process. In the same way as medicine, the research process is as much an art as a science and requires aesthetic skills. Applying Farquhar's (1994) position in relation to the practice of Chinese medicine I had to move from 'having knowledge of' research to 'knowing' research. Thus, it was in the doing of the research that I moved from an analytical understanding of research to an embodiment of the process.

Professional Development

The aim of the EdD is to help in the development of not just skills as a researcher but all aspects of one's professional life. As a teacher, the Foundations of Professionalism allowed me to pursue an area of interest on the teaching of reflective practice. Further, the lectures, seminars, assessment and feedback supported my development of the skills of structure, criticality and synthesis needed at doctorate level. The empirical research I conducted for MOE1, 2 and the IFS allowed for a more intimate relationship with the clinical reasoning literature and so my delivery of a course and feedback to students was steeped in a different scholarship. A key aim of my teaching of clinical reasoning was to develop practitioners as scholar-physicians by providing new insights into their practice and a language steeped in contemporary literature to explain it. As a practitioner of CM, the insights gained from the IFS have also improved my practice. In particular has been my recognition of the role of abduction that clarifies the place that CM places on a diagnosis being in that time and place. I have now presented this work in a number of situations – at professional conferences in the UK and internationally, at staff development workshops, and more recently at practitioner CPD groups throughout the country. I still need to find a tangible way to present the multi-modal aspect of the work that combined the film of a consultation with the interview of the practitioner. In terms of my role as an external examiner and accreditor, as said in my impact statement, I have already presented preliminary findings from my thesis research at a BAAB workshop and as a result I have been asked to be

the educational lead in the update of the Standards of Education and Training for my professional body. As an external examiner I now understand assessment differently and I provide more focused feedback to the managers and teachers in other institutions.

In conclusion I would like to propose that it is the dialectic between theory and practice that has to be understood both to practise medicine and to practise research and this is what the EdD supports. As Hamilton (2005) suggests, research when viewed as practice is emergent and each return to the data and the literature continues to illuminate different nuances and possibilities, and develop the multiple roles we hold as professionals.

Acknowledgements

I pay tribute to my supervisor, Will Gibson, who took me on in a time of need, and gave me the faith that what I had to say was valuable. His questioning of my sense making allowed me to develop a thesis that holds together as a coherent story. Caroline Pelletier has also given valuable feedback.

I thank the most generous contributions I had from the stakeholders in the research who gave me their passion and insight into the complex world of assessment. Colleagues I have journeyed with for years in education and in Chinese medicine education and accreditation have also been great sources of understanding and inspiration – Sibyl Coldham needs a special mention, and also Ann Rumpus. I thank the lecturers and fellow students on the EdD who each opened up ideas that have found a way into this study. And I must also thank the students and patients over the last 40 years who each in their unique way have allowed me to indulge in my passion for Chinese medicine.

Without the Hampstead Mixed Pond, which I discovered at an important point in this EdD journey, and the wonderful and watery companions I swim with, I might not have survived the journey.

This thesis is dedicated to my most loving husband who has been my rock throughout. The endless cups of tea, the proofing, wonderful discussions on the different meanings of words and his copy of Fowler's to sort out the use of -ize or -ise, have all been critical (or crucial).

Glossary

ACAOM   Accreditation Commission for Acupuncture and Oriental
            Medicine (USA)

ACGME   Accreditation Council for Graduate Medical Education (USA)

AHPRA   Australian Health Practitioner Regulation Agency

BAAB   British Acupuncture Accreditation Board

BAcC   British Acupuncture Council

CBE   Competency-Based Education

CBME   Competency-Based Medical Education

CM   Chinese Medicine

CMBA   Chinese Medical Board of Australia

NCCAOM   The National Certification Commission for Acupuncture and
            Oriental Medicine (USA)

WBA   Workplace-Based Assessment

WBL   Workplace-Based Learning

Table of Contents

List of Tables

Chapter One: Background

Introduction

> Do we choose to crash up against the rocky shoal of checklists and the
> atomization of medicine they promote or to be sucked down into the
> whirlpool that is subjectivity and the concerns about fairness and
> defensibility that go with it? (Eva and Hodges, 2012, p. 914).

In the dialogue that takes place between Eva and Hodges, the question they
are posing concerns the navigation between objectified checklists and
subjective judgement in the process of the assessment of students'
competence to practise medicine. Their debate continues, that while we, as
medical educators, are critical of objectivity, we demand systematic
measurement and while we yearn for the recognition of an assessor's personal
and holistic[1] judgement, we are wary of its 'inefficiencies and lack of uniformity'.
The two words, objectivity and judgement, Eva and Hodges discuss, are not
neutral, they can mean different things depending on context and present
different ways of understanding the world. The demand for standardization in
the form of competencies and measurement is now embedded in Higher
Education (HE) institutions but, so too, I argue, is subjective judgement
(Harman and McDowell, 2011), held within the role of the supervisor as
assessor, more implicit perhaps and less researched.

This thesis explores how educators in three institutions teaching Chinese
medicine (CM) in different English speaking countries have dealt with the
tensions introduced above as competency-based medical education (CBME),
as a curriculum model, is being globalized. Hodges and Lingard (2012a) warn
that teaching and assessment methods are not spontaneous but are adopted
by institutions - sometimes unconsciously so - according to particular and
dominant discourses of pedagogy, curriculum policies and external
stakeholders. My interest is what form workplace-based assessment (WBA)

---

[1] Holism is a term that denotes plurality, complexity and a socio-constructive approach to
teaching and assessment, used within education in contrast to an analytic or reductive
approach (Bleakley, 2010; Govaerts, van der Vleuten and Holmboe, 2019). The 'holism' of CM
will be explained on page 87.

was taking in institutions teaching CM and how the constructs of CBME were being experienced.

1.1 My context

In order to create context for this thesis I will present my personal relationship to assessment within CM education and then present a brief background on the place of CM education within China and within the UK. Standardization and institutionalization form a theme throughout this section.

I was one of the founders and have been a director of a Chinese Medicine professional-entry training course, originally in a private school set up in 1983 by a group of practitioners trained in the UK and in China. In 1996 the course was validated as a BSc (Hons) Chinese Medicine: Acupuncture degree and moved into a UK University where I was the course leader and principal lecturer for 23 years. Over my years of being immersed in the teaching of CM I have been involved in every aspect of running a school and course. I have written curricula and codes of practice; interviewed and recruited students and teachers; decided on clinic layouts and sourced equipment. I have planned clinic schedules and inducted practitioners into clinical teaching. I have interpreted competencies and prided myself on writing carefully constructed and integrated learning outcomes and criteria for clinical assessment instruments; I have been a clinical supervisor teaching and assessing students applying those instruments and have moderated other supervisors' assessments. I have been an external examiner considering whether assessments meet quality assurance requirements and as an accreditation officer for the British Acupuncture Accreditation Board (BAAB) determining if other courses meet professional body standards. I was part of a working group that produced the Standards for Education and Training for Acupuncture (SETA) of the British Acupuncture Council (BAcC, 2011) and subsequently involved in mapping competencies onto the University degree course. Having been intimately involved in all aspects of clinical training, my interest has turned to look more closely at how we assess students' work with patients, especially the application of learning

outcomes and assessment instruments to determine students' competence to practise. My involvement in all the parts that form the 'architecture' of a course (Kemmis and Mahon, 2017), as described above, have led to this thesis.

I found myself asking a clinical supervisor the following question one day when she was defending her assessment of a student at the end of his placement in the teaching clinic.

> *"But is he competent?" I ask the supervisor. "Are you happy for him to go out into practice and treat patients? Would you send someone to him for treatment?"*

As I reflected on this conversation I noticed how my questions to the supervisor were appealing to her role as a practitioner of CM. It was not about where the student was scoring on the competencies checklist or which sub-competencies he was failing in; it was not about what final mark he was to be given. The question was whether, in her professional judgement as a practitioner of CM, she thought this student was safe to be treating patients, safe to join our community of practice. I noted a discord in my approach to assessment. In order to try to support clinical supervisors in their role, my tendency had been to be more and more explicit in the learning outcomes and criteria within the assessment instrument. However, the instrumentalism of learning outcomes that has become more dominant in HE and more embedded over the last 10 years of university periodic reviews, did not seem to be supporting the process of WBA for supervisors. My interest turned to what part the role of judgement played in assessment. How was the assessment instrument aligning with the judgements of the supervisors? How might a lack of objectivity lead to errors of measurement? And, what might be impacting on our role as supervisor-assessors? These are the questions behind my decision to carry out this research.

1.2 Why do we assess students?

The need to provide a grade of a student's competence to practise is a requirement of both the professional bodies and Universities. In the UK the BAAB (2016) states:

SETAP 6.5 The <u>measurement</u> of student <u>performance</u> must be <u>objective</u> and ensure fitness to practise (2016, p. 45).

The expressions of measurement, performance and objectivity form a substrate to this thesis. How do we measure a student's performance in a clinical setting, an assessment that is conducted simultaneously over time and in the moment? This thesis aimed to find out what form objectivity was taking in WBA.

Assessment in any institution has multiple purposes. It must be summative and formative, efficient and effective and meet the needs of multiple stakeholders (Govaerts, van der Vleuten and Holmboe, 2019). It provides data to the institute for quality assurance; it indicates to teachers their success (or otherwise); it promotes (or should promote) learning; it helps inform the student of their strengths and weaknesses and gives them the assurance to go out into the world; it certifies achievement for the student and any regulatory body or employer (Boud and Falchikov, 2007; Shumway and Harden, 2003; Trede and Smith, 2012; Yorke, 2005; 2009). In the case of assessment for professional entry it also informs the public and the practice community of a person's legitimacy and safety (Trede and Smith, 2012). This thesis aimed to investigate the different meanings that these purposes created for the different stakeholders.

1.3 History of Chinese Medicine Courses

In order to contextualize the place of CM teaching institutions in the Western[2]/Anglo-American world this section will provide a brief history of Chinese medicine from its place within the politics and culture of China. This will provide the wider perspective on how standardization and institutionalization

---

[2] The use of the expression 'West' and 'East' in opposition, more an expression of cultural than geographic difference, is an artificial construct that, in relation to medicine, can misrepresent the hybridity and heterogeneity that exists globally. I propose, however, that it is valid in relation to a thesis about Chinese medicine and it's practice in the Western/Anglo-American world. I use the expression 'Western' medicine to depict what is variously known as allopathic, orthodox, conventional or bio-medicine and, in the UK, state sponsored medicine. Likewise the expression Chinese to encompass the East has also been adopted as the institutions that have been researched are teaching a predominantly Chinese medicine as opposed to Japanese or Korean medicines, which, while historically entangled, are seen as distinct forms.

have impacted on CM education. I will then move to the 20th/21st centuries and the situation of CM in the UK.

Scheid's (2002) analysis of the history of medicine in China shows it to be one of heterogeneity and multiplicity, permeable to outside influences of time and place, culture, politics and technologies. For most of the last 2,000 years, the scholar-physicians learned primarily through a lineage system in which the student followed a master studying the classical canons and at the same time gaining their experience of how to gather, synthesise and analyse information from patients through their skills in observing, questioning, palpating, listening and smelling. Over the last two centuries, processes of institutionalization have impacted on the practice of medicine in China and the training of its doctors, each time creating something different for the stakeholder, be they the state, the doctor or the patient. The very concept of a 'Chinese' medicine only came into existence when doctors were forced to define their medicine as different from that of the West (Andrews, 2015; Taylor, 2004b), no longer '*yi'* (medicine) but now *zhongyi* (Chinese medicine) as different from *xiyi (*Western medicine)*. It is important to note that this is a medicine that has a different philosophical, cultural and scientific approach*, one in which the cartesian split of object and subject has not been a part (Rošker, 2018). However, since 1949, the aim of the Communist party in the PRC has been an integrated *zhongyi* and *xiyi*.

In 1956 as part of China's modernization, a new form of college of Chinese medicine was founded (Scheid and Lei, 2014), and, while not engulfed by them, the institutional, clinical and even pedagogical standards of Western medicine now operated (Zhan, 2009). Lei (2002) describes, for example, how doctors trained by masters and not in the universities had to sit examinations in *xueshu* (canonical learning) first and only if they passed would they then be invited to take the *jingyan*[3] (experience) test. If a doctor had only studied with a master and not in a university they needed to take supplementary classes. The studying of books in other words, took precedence over experience thus introducing a rupture from the traditional focus of CM as a practice (Scheid,

---

[3] Lei (2002) explains the ramifications of jingyan being 'incorrectly translated' as experience, but that nuance does not impact on the meaning I am using here.

2016). Farquhar (1987, p. 1020) explains that while the canonical medical texts of the Han dynasty such as the *Huang Di Nei Jing* might 'exercise a powerful constraint over the knowledge claims that can be made [by physicians]', they are also contingent on professional practice as 'their virtue lies precisely in their adaptability to the uses of the present'.

In the 1960's the Chinese Ministries of Health and Education commissioned a series of textbooks. The aim was to unite the many competing currents within CM and the contradictory information in the classical canons and at the same time to translate the canons from classical into modern Chinese (Andrews, 2015; Scheid and Lei, 2014). In the PRC's drive for scientization and modernization, CM had to rewrite its canons showing a 'Western' systematized biological logic that down-graded the information of the senses and the unity of body-mind-spirit (Barnes, 2003). According to Taylor (2004a) this was not just another example of standardization that had occurred at various stages in the 2000 year history of CM but, she proposes, this petrifying of CM knowledge into a standardized form changed the transmission of CM by fundamentally altering the dynamic of master-disciple, of patient-doctor. The 'dialectic between clinical insight and scholarly learning', one of its defining features, Scheid (2008, p. 487) claims, was substantially lost.

As the medicine of China has been moved, translated, interpreted and translocated to the West and into HE institutions with their own standardizations, it has taken on new forms (Zhan, 2009). The first training courses were established in the UK in the 1960's and 70's, with a proliferation of CM schools occurring through the 1980's and 90's (Uddin, 2008). These private schools were set up and run by practitioners of CM trained in the West and in the PRC. Clinical training was either through observation in private clinics or clinics set up by the schools. The supervisors were all practitioners teaching part-time and concurrently managing their own private practices and mostly with no teacher training.

As part of the professionalization of CM, professional bodies were formed introducing standards, rules and regulations on ethics and practice and

accountability processes for patient safety (Barnes, 2003; Flesch, 2013). Accreditation procedures for teaching institutes were also introduced and in the UK the BAAB was formed in 1990. In order to manifest appropriate (Western) educational standards, schools adopted HE structures and processes and from 1996 some became externally validated or embedded within Universities. At its height in 2007 there were 10 courses in the UK accredited by the BAAB, with 5 of them located within universities (Uddin, 2008). This has inevitably had an impact on the way CM has been taught, assessed and practised (Givati and Hatton, 2015). Barnes (2003), Cant and Sharma (1999) and Flesch (2013) suggest that, as in the PRC, it is the very alignment with Western medical models of professionalization, education and practice that has led to the increased acceptance of Chinese medicine in the West. However, this dominance of a Western medical model, they explain, also brings problems. Within the history and translocation of CM to the West, another of my questions for this thesis was how standardization and institutionalization were influencing WBA.

1.4 Research Questions

In this thesis the context presented above is given as background to this investigation.

> How is competence to practise being assessed in programmes of study that lead to professional entry in Chinese Medicine?

By investigating three CM institutions in different countries, my aim was to illuminate their chosen theories of assessment. This understanding, I propose, could help accreditors, institutions and supervisors to develop more appropriate assessment processes. The following 3 sub-questions directed the research.

1. What are the similarities and differences in WBA between the three accreditors studied?

2. What are the similarities and differences in WBA within and between the three institutions studied?

3. How is CBME constructed by the different stakeholders (accreditors, institutions, managers and supervisors), and what issues does this present?

My original aim was to investigate CBME and the tensions between the instrumental-technical and socio-constructive perspectives on WBA. The fourth question was derived after I had started the data analysis and realized that Schatzki's practice theory could be applied to assessment to illuminate the social, cultural and material spaces in which WBA resided.

4. What are the social-cultural-material activities and arrangements that constitute the practice of assessment in the three institutions studied?

1.5 Structure of thesis

This thesis comprises 8 chapters. Given that the accreditation bodies in the three countries I have focused on are promoting a CBME curriculum model, Chapter 2 reviews the literature on CBME in relation to WBA. While the assessment of clinical work in the biomedical health professions has been much debated and researched, for example, Govaerts *et al.* (2013), Higgs *et al.* (2012), Holmboe *et al.* (2017), Schuwirth and van der Vleuten (2014) there has been no research into what form this takes within CM courses. This reflects, I propose, a profession in the early stages of its development in the West. This review, therefore, considers the empirical and theoretical literature within the mainstream health professions to form a framework to analyse WBA in CM. This chapter also presents an alternative conceptual framework to that within CBME with which to analyse WBA, namely Theodore Schatzki's 'practice turn', which considers a phenomenon from a socio-cultural-material position. Using case study as my research method and a socio-interpretivist approach, I explain in Chapter 3 the process of investigation including the opportunities and challenges of being an insider/outsider researcher. I describe the thematic

analytical process used as I constructed meaning and the ethical considerations that permeated the research process. Chapters 4-6, the analysis and discussion chapters, are constructed around the stories of the three main stakeholders – the accreditors, the institutions (including the managers) and the supervisors. Through focusing on the similarities and differences between the different institutions, these chapters explore their experiences of WBA and the tensions created by the adoption of CBME. Chapter 7 applies my alternative theoretical approach to the data. By using the framework of assessment as a 'practice', different explanations of the enactment of assessment and the entanglement of the different stories being told by the stakeholders, are foregrounded, stories that are not being considered within the CBME literature. Chapter 8 will draw the thesis together and consider the main issues that the thesis presents for the development of WBA in CM institutions in the West and for WBA in general. A reflection on the limitations of the study will help explain future directions for research.

Chapter Two: Workplace-Based Assessment

Introduction

In this chapter I introduce the unique features of learning in the workplace and give a brief summary of the origins and applications of CBME to show some of the epistemological positions and filters that suffuse the literature on workplace-based assessment (WBA)[4]. The tensions in CBME, which were illustrated in the discussion between Eva and Hodges in the introduction, can be viewed as polarities (Govaerts, van der Vleuten and Holmboe, 2019), a dialectic (Kemmis, 2009) between the two theoretical approaches of the instrumental-technical and socio-constructive. Taking this position in my review of the literature on WBA was congruent with the multiple roles I have had in relation to WBA. In section 2.2 I expand on two main constructs of CBME arguing that there are clear tensions between standardization and authenticity and between atomization and holism. In 2.3 the focus is on measurement, the role of psychometrics and the place of validity and reliability and the construct of a single true score in the assessment of a complex performance that is context specific. In 2.4 my attention turns to the role of the assessor and views on bias and judgement. In order to theorise WBA further, Section 2.5 introduces Schatzki's (2001; 2012) 'practice turn' as an alternative theoretical approach to the instrumental-technical/socio-constructive approach to WBA.

2.1 Learning in the workplace

The necessity of including workplace experiences, and often for quite extended periods of time, has long been acknowledged within the major professions, disciplines and trades e.g. doctors, engineers, carpenters (Billett, 2010; Hager, 2011; Higgs, 2012b; Yorke, 2005). According to Evans, Guile and Harris (2011) it was in the late 80's and 90's that workplace-based learning (WBL) was

---

[4] I have used the expression 'workplace-based' instead of 'practice-based' throughout the thesis, holding the word 'practice' for the particular approach taken by Schatzki as in the 'practice turn' and presented in Chapter 2.5. There are distinctions behind the terms workplace-based and practice-based and I refer the reader to Evans and Guile (2012b) for clarity. For the purposes of this thesis those distinctions are not relevant.

introduced into institutes of higher education (HE) in the UK, not just for occupationally specific programmes of training but learning in general. Experiences gained in workplaces were seen as essential for developing the sort of knowledge required for 'work-ready' graduates (Allan *et al.*, 2011; Caballero, Walker and Fuller-Tyszkiewicz, 2011; McEwen and Trede, 2014) and employers and professional bodies became central to decisions on the outcomes of the curriculum (Bagnall and Hodge, 2017; Morcke, Dornan and Eika, 2013). Schwandt (2005) proposes that the bringing of the new health professions into universities, with placements located in hospitals or clinics, has enhanced their professional status. This move into HE, however, has brought with it new standardizations and processes of accountability for the practitioners, educators, academics, teachers and students involved, including for CM courses.

The learning that takes place in the workplace has a different and distinct pedagogy Kemmis (2012) argues, a different 'logic' (Evans and Guile, 2012b), from classroom or didactic forms of learning with the workplace needing to be seen as an important and different learning environment in its own right. Young (2008) proposes that a curriculum must provide access to the canonical and fixed knowledge of the discipline but also the contextual, situationally determined and tacit knowledge that comes from the workplace and not attempt to collapse the distinctions between the two. It is in the workplace, as Scheid (2008) has argued, that the dialectic between canonical knowledge and experiential knowledge (Blinman, 2017; Eva and Hodges, 2012; Rotthoff, 2018) along with the socialization into the profession, is manifest. A concern is raised by Schwandt (2005) that the subjective, interpretive, inconsistent and intuitive nature of professional work that is required in the work-place is often 'remedied' by the application of canonical[5] knowledge or what Orr (2007) calls 'scientific' knowledge.

---

[5] Canonical: also propositional or theoretical or codified. I use canonical as it signifies the 'canons' that are the basis of CM and because Billett's (2017) application of the word includes procedures as well as concepts.

2.1.1 Competency-based Medical Education

It was to bridge these different worlds of classroom-based and work-based education and the tension between the learning and assessment of canonical knowledge and the professional knowledge needed in the workplace, that an outcomes-based approach to curriculum design was introduced into HE. It was also the introduction of more neo-liberal political and economic systems into Universities, the development of a performative audit culture and quality assurance frameworks that led to the focus of education in HE to shift (Grant, 2014; Hodges, 2012; Preston, 2017; Stone, Boud and Hager, 2011). The curriculum was no longer to be based on the traditional 'input' perspective of the process of teaching a student a certain content within a certain time frame (Carraccio *et al.*, 2002; Fishbain, Danon and Nissanholz-Gannot, 2019; Harden and Crosby, 1999) or the standards of the masters embedded in the guilds (Kemmis, 2008), but now constructed around 'outcome' measures, the measurement of students' performances, the product of students' achievements. Pre-defined and pre-specified competency statements would act as a frame of reference for institutions to set learning outcomes and assessment criteria in order to support the decision making of assessors, provide the basis for an objective marking system and make the requirements for successful completion of the course explicit for students. Referred to as competency-based education (CBE)[6], this curricular model was manifest within my own University, has been assumed in the development of the Standards for Education and Training of Acupuncture (BAcC, 2011), and was seen in all the accreditors I investigated

Within medical education defined competencies that combine the knowledge, skills and attitudes expected of physicians have been introduced for some 20 years (Holmboe, Edgar and Hamstra, 2016; Park, Hodges and Tekian, 2016).

---

[6] Much of the literature including Kennedy's AMEE Guide of 2014 uses the two terms competencies and outcomes interchangeably. I have tended to use competencies when referring to a general concept and when it is used as CBE/CBME and outcomes for what is described within assessment instruments. Any direct quotes however have followed the author's original word choice.

The AMEE guide by Smith and Dollase (1999) states:

> We believe that by clearly specifying the educational outcomes in behaviorally measurable ways, we can change the way faculty teach and students learn. (p.15)

The following definition of CBME appears on the International Competency Based Medical Education (ICBME, 2018) Collaborators website:

> CBME is a method of assuring the production of competent physicians by utilizing explicit abilities (or competencies) of physicians and using these competencies as a way to organize medical education.

These two statements illustrate that CBME was to be an approach that would organize not just the curriculum but also the pedagogy.

Dr Jason Frank, the Co-chair of the ICBME Collaborators and others, in their definition of competency, present another key aim of CBME.

> Competency: an observable ability of a health professional related to a specific activity that integrates knowledge, skills, values, and attitudes. Since competencies are observable, they can be measured and assessed to ensure their acquisition. Competencies can be assembled like building blocks to facilitate progressive development. (Frank *et al.*, 2010b, p. 641)

This definition reinforces how the assessment of competence would now come about through measuring the observable behaviours of students' performances rather than defining the process of education. The type of doctor to be produced (Harden and Crosby, 1999) would be defined through competency statements that incorporated the domains, tasks and roles within a specific performance considered essential for the practice of medicine. By the application of appropriate assessment instruments to ensure and measure if students had attained those competencies, the accountability requirement of society would be met (Frank *et al.*, 2010a; Hodges, 2010).

CBME has been embraced wholeheartedly internationally (Hodges and Lingard, 2012b; ten Cate, 2017). It can be seen in a variety of medical education accreditation systems for example the General Medical Council (GMC, 2018) in the UK and the Accreditation Council for General Medical Education (ACGME, 2017) in the USA. The key drive has been to 'standardize learning outcomes

and general competencies' rather than to focus on the 'length and structure of the curriculum' (Irby, Cooke and O'Brien, 2010, p. 224). This dominance of a competency approach in medical education is in spite of there being little empirical evidence that it achieves better results than previous training systems (Morcke, Dornan and Eika, 2013). While this lack of research into the effectiveness of CBME was acknowledged in the paper by Holmboe *et al.* (2017) who are members of the ICBME Collaborators, it is still widely advocated.

From the perspective of this thesis it has been important to treat definitions of 'competence' not as value free, but as social constructs, politically and ideologically motivated (Rotthoff, 2018; Yorke, 2005). For Hodges (2012) the language of competence is not just descriptive but also constructive. He identifies a number of discourses which are enmeshed and recur in the literature on CBME and which influence the concept, often implicitly. Competence can be seen as knowledge, as performance, as production and as psychometrics, each of these impacting on how assessment of competence can be understood, portrayed and conducted. The knowledge discourse of competence manifests in the requirement for the memorization and reproduction of information with incompetence manifesting as a lack of recall of facts and concepts or incorrect following of routine procedures. Competence as performance is to do with observable behaviours as opposed to cognitive processes and incompetence is usually interpreted as an absence of the required behaviours. Those behaviours are often used as 'proxy' (ibid. p.25) indicators of underlying attributes or abilities. Competence as product, as outcomes, Hodges (2012) relates to a quality assurance and accountability discourse with students needing to conform to given standards. In the psychometric or measurement discourse, human characteristics are represented by numbers, and incompetence manifests as an inability to score a pass on the standardized measure. Hodges recommends a sensitivity to these discourses as important when considering the literature on WBA.

There are many critiques of CBME some of which are discussed in more detail in the next sections. Given my interest in the application of the literature to

Chinese medicine I note a concern from a number of educationalists around issues of (Western) cultural dominance within CBME. Ho *et al.* (2011) report differences across cultures (Eastern and Western) especially in the interpretation of competencies around professionalism and self-integrity; Ho *et al.* (2012) reveal the prominence of Confucian relationalism impacting on descriptions of competence in Taiwan; and Ho *et al.* (2014) discuss how social and personal roles are differently integrated to those in a Western curriculum. Bleakley, Brice and Bligh (2008) and Grant (2014) highlight the influence of cultural and historical factors on the very concept of competencies, showing evidence of different interpretations in the language of competencies across countries. While a universal definition of competencies has been attempted in medicine, McGaghie *et al.* in 1978 were warning against this due to the differing political, social and cultural circumstances of institutions and courses. Albanese *et al.* (2008) insist that the key criteria for determining competencies is that they should reflect the vision of the particular institution as well as any stakeholders including the public, leading to tension as CBME is being globalised. Researchers in the PRC, as example Wu *et al.* (2015, p. 4), are calling for CM theory to be written in a scientific language rather than philosophical or even empirical, in order to make it 'universally accepted by the modern scientific community'. Barnes (2003) is concerned that Western medicine is becoming the prototype against which all other systems of healing are being measured, the arbiter of what is even considered 'medicine' (Saks, 2015), and that includes professional competencies. That Eraut (1994) is concerned that the political purpose of CBME is overriding its educational value reminds us of the origins of CBE and its birth in an instrumental, technical and behavioural milieu in a particular place (the West) and time.

2.1.2 Assessment of Workplace-based Learning

The direct observation of the clinical encounter in the workplace that focuses on how students integrate skills, knowledge, judgment and attitudes is proposed by van der Vleuten *et al.* (2010) as both more authentic, a more appropriate stimulus format and pitched at the right level of complexity for assessing professional competence. However, considering's Frank *et al's* (2010b) definition of competence as described earlier, and taking Epstein and Hundert's

description:

> The habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, judgement, emotions, values and reflections in daily practice for the benefit of the individual and the community being served. (2002, p. 226)

then we see a number of possible tensions. Assessment of competence will need to be made of not just skills and knowledge, but also the affective state of the student and their clinical reasoning, an invisible process that is often hard to define albeit manifest in behaviours or artefacts. At the same time de Jonge *et al.* (2017) remind us of the complexities and uncertainties and uncontrolled variables of patient mix, numbers and case difficulty in the workplace. While the defining of competencies in medical education might be seen as straightforward, what challenges institutions following the CBME model, is how to translate professional performance into a standardized list of outcomes against which to measure students' competence to practise (Grealish, 2015; Tekian *et al.*, 2015).

2.2 Rocky shoals and whirlpools

Reminded of Eva and Hodges discussions in the introduction on the place of systematic measurement and holistic judgement, this section will present the literature on CBME under the tensions between standardization and authenticity and between atomization and holism that emerge from the different constructs inherent in CBME, namely the instrumental-technical and socio-constructive.

2.2.1 Standardization and Authenticity.

> WBAs can evaluate multiple, essential competencies simultaneously in an integrated fashion in the authenticity of day-to-day practice. (Kogan and Holmboe, 2013, p. S68)

This statement holds within it one of the main tensions within CBME that is, the standardization of competencies and the authenticity of the setting of students' learning (Govaerts, van der Vleuten and Holmboe, 2019). Proitz (2010, p. 122) describes the argument in this tension as a continuum between a behaviourist and social-constructivist approach. In the former, competencies are 'result-

oriented, full-ended and measurable' and educational interventions are seen as a cause that will have an observable measurable effect that can be quantified pre- and post-test. The latter sees competencies as 'process-oriented, open-ended and with limited measurability'. Authentic performance cannot be pre-ordained as it is the locus of learning which creates and shapes the opportunities and affordances for developing the knowledge and skills needed in competence to practice (Rees, 2004; Villarroel *et al.*, 2018). The instrumentality of CBME is manifest through fixed and explicit competency statements, assessment instruments and collection of metrics along with a focus on performance. At the same time the socio-constructivist discourse of CBME calls for learning to be rooted in authentic settings that recognizes the contextualized nature of knowledge, and that assessment will need to capture a student's day-to-day performance as they are immersed in patient care. Within WBA is a tension between fixed outcomes and contextualized knowledge.

It is the level of codification that determines how easy or difficult different forms of knowledge are to contextualize as competencies (Wheelahan, 2009). Canonical knowledge is easier, practice knowledge more difficult, while being uncodified and tacit, personal forms of knowledge provide few principles on which to base competencies. Thus standardized competencies are limited to knowledge, skills and aptitudes that can be manifest as overt behaviours, while what needs to be assessed is a complex set of concepts, procedures and dispositions, an 'engaged, embodied and enacted judgement that links knowledge, virtue and reason' (Schwandt, 2005, p. 321).

Rethans *et al.* (2002) present a differentiation that sheds light on these tensions but is rarely made in the literature. It stems from the quite different concepts of knowledge as described. They define competency-based assessments as 'measures of what doctors do in testing situations', and performance-based 'as measures of what doctors do in practice' (p. 901). While pre-specified competencies might lend themselves to the assessment of factual knowledge and skills, or what practitioners do in controlled conditions such as in objective structured clinical examinations (OSCE's), the defining and 'measurement' of human behaviours in the workplace is impacted on by the lack of certainty and

standardization and the complexity that comes from the very authenticity of that place and the multiple dimensions that need to be assessed. When knowledge was the dominant view of competence it was the poor integration of that knowledge with appropriate interpersonal skills that was a major concern in medical education (Morcke, Dornan and Eika, 2013). However, Barnett (1994), Hyland (2017), Talbot (2004) and Wear (2008) all convey a concern that the logic behind competencies in the domain of skills and knowledge which can be specified as behaviours, has been illogically imposed on areas such as the context-driven nature of empathy, humility, respect and a deeper reflective and humanistic engagement with the practice of medicine. Cooke, Irby and O'Brien (2010) in their Carnegie Foundation Report consider that defining and assessing 'excellence, humanism, accountability and altruism' would be 'difficult to reconcile with a competency-based framework' (p108). If competency statements need to describe and embed not just non-technical or non-cognitive skills but also professional and personal values such as compassion or caring, the problem comes with defining the latter as explicit and objective statements which being elusive to observation are difficult to quantify.

The standardization of competencies for purposes of producing objective measurements in terms of validity and reliability has been a major factor in the development of CBME (ICBME, 2018). At the same time, the context in which the assessment is being conducted is equally important. The application of standardized competencies will struggle to capture contextualized and emergent learning in the workplace (Hager, 2011). Thus WBA may not be as easily managed as stated by Kogan and Holmboe (2013) at the beginning of this section.

2.2.2 Atomization and Holism

> The measurable bits of performance that follow from anatomizing clinical competence according to discrete learning objectives do not and cannot add back together to constitute the skill and ability of the competent physician. (Huddle and Heudebert, 2007, p. 537)

This statement sums up a further tension in CBME, that between atomization and holism (Ginsburg *et al.*, 2010). The aims when competencies are broken

down into sub-competencies that contain more specific performance instructions are to: clarify and ensure that all the behaviours to be manifest are taught and assessed (Govaerts, van der Vleuten and Holmboe, 2019); ensure that student deficiencies in one domain are not compensated by another (Lucey, Thibault and ten Cate, 2018); allow for fairness and equivalence in assessments; and reduce assessor bias and different performance expectations (Crossley and Jolly, 2012). Holmboe *et al.* (2017) add another important feature. Discrete sub-competencies allow for the 'granularity' needed in feedback to students so that the reasons for failure can be clearly identified. Further, they argue, 'the degree of acceptable variation in assessment should be "bounded"'(p.576) and that is what the sub-competencies provide. As Hager (2004) notes, atomistic learning can be attractive as it offers a simple approach to assessment with direct observation of work done being checked against a list. This product view of competence, however, comes with its own problems.

It is not the specified general domains, the overarching competencies, that creates the problem in CBME, but that in the breaking down of performance into discrete, multiple and explicit outcomes, the complexity of what professional competence involves is denied. It is the 'an-atomizing' of performance, what Kuper, Whitehead and Hodges (2013) call the dissection of the 'good' doctor into their component parts, that is challenged as it neglects the complexity of the interrelatedness of knowledge and understanding required in professional work (Blinman, 2017; Ginsburg *et al.*, 2010; Orrell and Higgs, 2012; Regehr *et al.*, 2007). While competency frameworks deconstruct, it is the assessors who need to reconstruct these atomized statements back together and consider the whole performance. For ten Cate and Billett (2014) learning outcomes, as 'abstractions', do not always link clearly to the tasks within the workplace, making it difficult for assessors to reconcile what they are observing with the published statements. According to Crossley (2013, p. 1153) atomizing, as a drive for objective assessment, has had the 'paradoxical effect of increasing the variability between judges rather than reducing it'. Research also shows that the more sub-competencies there are, or the longer the check-lists, the more chance there is of the assessor reverting to an holistic assessment (Kogan *et al.*, 2015). A systematic review by Lurie, Mooney and Lyness (2009) shows that

competencies are not independent constructs but overlap and are interdependent and so difficult to measure as discrete items. It would seem that the holistic manifestation of competence cannot be described with the objective precision warranted by CBME.

What competencies describe, Wheelahan (2009) contends, is just those things that can be observed or measured neglecting the more higher-order thinking, attitudes and values that cannot. The conclusion is that this defines what is assessed to the limits of language. Tummons (2020) critiques the conflation of competence with competency statements as one of mistaking actual things in the world with the ways in which we write them. One of the advantages proposed by atomized competencies is that they can foster transparency and consistency of interpretation by the professional community – accreditors, teachers, assessors, students - 'shared mental models of what needs to be achieved' (Govaerts, van der Vleuten and Holmboe, 2019, p. 70). Bloxham, Boyd and Orr (2011) and Kemmis (2008), however, describe how assessors report the terminology in competencies as often being ambiguous and confusing with multiple possible interpretations, and therefore resort to their own personal and more holistic understanding of competent performance (Almalkawi, Jester and Terry, 2018; Govaerts *et al.*, 2013).

There is a different view on the place of the atomization of competencies that loosens their link directly from the measurement aspect of assessment. Tekian *et al.* (2015) recommend that the sub-competencies should be seen not as a list of activities to check off but more as a 'dictionary' to narrate competence, or as Eva and Hodges (2012) propose, competencies should be seen as a 'guide' only and we should let go of the 'notion that we can measure competence' (p.918).

Whitehead *et al.* (2015) sum up the situation of the standardization/atomization and authentic/holism debate:

> competency frameworks are extremely valuable in that they provide a
> clear and convenient, if somewhat abstract, model of complex real world
> systems. As abstractions, however, they necessarily provide a simplified
> version of the complex ideas they represent and the ways that practice

changes in response to the context in which it takes place. No framework, therefore, is ever "the truth", but instead all frameworks are approximations, and all will inevitably have limitations as well as strengths. (p.246)

Maintaining an awareness of the different assumptions that arise when WBA is seen in the context of standardization and atomized statements of competence or of holism that needs to be captured when learning is in an authentic setting, will allow for a more informed dialogue in curriculum and assessment development. Even Holmboe *et al.* (2017) present a different approach from their 2013 statement as they have seen the tensions that arise from the different epistemological positions of the instrumental-technical and socio-constructive.

Holism and reductionism both have a place in medical education; the issue is when and how to apply these interconnected philosophies appropriately. (p.576)

Developing these 'philosophies' further the next section will look at the role of measurement in WBA. The dissection of competence into multiple sub-competencies, as discussed previously, is intimately tied to a measurement discourse.

2.3 Measuring Performance

When the public is concerned about the educational productivity of its schools the tendency, and it is a strong one, is to tighten up, to mandate, to measure, and to manage. (Eisner, 2004, p. 3)

To support the accountability role of accreditors and institutions the approach of CBME is to measure a student's performance against the competency statements developed by the particular health profession. When measurement is central to the process of assessment, the psychometric discourse can take a prominent position. This in turn raises issues of the validity and reliability of assessment instruments and the tensions of assessing in uncertain and uncontrolled authentic settings.

2.3.1 Psychometrics

The conversion of human behaviours to numbers through the objective measurement of knowledge, skills, abilities and attitudes is referred to as psychometrics. According to Park, Hodges and Tekian (2016), the number as a psychometrically reliable test score has been the backbone of assessment standards in professions as it provides the evidence to stakeholders that the assessments are providing a rigorous and efficient testing of competence, a critical issue in supporting the status of any profession.

At the heart of the psychometric discourse, however, are a number of assumptions. Firstly, that the acquisition and performance of discrete sub-competencies that are easily described and measured equals competence. Secondly, professional competencies as concepts, procedures and dispositions can easily map onto pre-determined and atomized outcomes (Govaerts and van der Vleuten, 2013; Morcke, Dornan and Eika, 2013). Thirdly, human characteristics and behaviours can be represented by standardized scales (Delandshere and Petrosky, 1998; Hodges, 2013). And finally, competence as knowledge and skills is an individual possession, stable and context free (Lingard, 2012; Schuwirth and van der Vleuten, 2006). The striving for validity and reliability and the concept of the 'single true score' are also framed within this discourse (Hodges, 2012; van der Vleuten *et al.*, 2010; Wass and Archer, 2011).

2.3.2 Validity and Reliability

For Shumway and Harden (2003) validity is concerned with whether the scores from an assessment instrument measure what they are supposed to measure and can distinguish levels of achievement between different students. Reliability is whether the assessment instrument consistently measures what it sets out to measure when repeated by the same rater, by different raters and over time. Hodges (2013, p. 565) adds a slightly different nuance describing reliability in authentic contexts as 'the accuracy of the transformation from real-world phenomenon to a number' and validity as 'the stability of that numerical representation'.

The improvement of assessment instruments for objectivity, fairness and transparency has been a major factor in the development of CBME (Crossley *et al.*, 2011; Holmboe *et al.*, 2010; van der Vleuten *et al.*, 2010; Wass and Archer, 2011). For students, grades hold both extrinsic and intrinsic value thus grade integrity is an issue for any institution (Sadler, 2010). In order to minimize inconsistencies in marking and reduce the bias of assessors, learning outcomes are advised to be explicit and standardized, assessment instruments valid and reliable, practical and efficient, and 'decisions on grades should be made by evaluating the quality of student work against fixed anchor points' (ibid, p.728). Examiners should have clear guidelines for marking, 'which indicate how performance against targeted curricular outcomes should be rewarded' (Shumway and Harden, 2003, p. 570).

While direct observations of the clinical encounter in an authentic setting as an appropriate assessment of a student's performance, and its representation in numerical form, have been the mainstay of clinical assessment in health care professions, there is much critique of their lack of validity and reliability (Blinman, 2017; Morcke, Dornan and Eika, 2013; Norman, Norcini and Bordage, 2014). Further, there is very little empirical evidence to support validity (Natesan *et al.*, 2018; Shorey *et al.*, 2019) or inter-rater reliability (Klamen *et al.*, 2016; Rencic *et al.*, 2016) with rater bias also being an issue (Sherbino *et al.*, 2013; Tavares, 2014). Holmboe *et al.* (2017), while supporting CBME, acknowledge that the summative assessment of competencies as applied in the workplace lacks evidence of validity and reliability. What Delandshere and Petrosky (1998) question is the use of numbers as a representation of what is a complex performance.

The conflation of the numerical with objectivity, reliability and fairness has been a major question in CBME for some time with Schuwirth and van der Vleuten (2019; 2006) providing in depth critiques. Their argument starts with how the term 'subjective' has been framed in opposition to 'objective' with the consequence that in relation to assessment subjective has come to mean biased and biased with unfairness. What they conclude, instead, is that

objectivity and reliability do not necessarily equate, and conversely subjectivity does not necessarily relate to unreliability and hence unfairness. An objective, standardized check-list of performance can produce unreliable scores while subjective, global rating scales based on expert (and subjective) judgements can provide reliable scores and might be better able to discriminate between levels of expertise. To improve reliability and validity extended sampling with multiple raters assessing multiple encounters in multiple contexts is needed (Lockyer *et al.*, 2017).

Rather than even considering expressions such as validity and reliability and objectivity Govaerts and van der Vleuten (2013) recommend drawing inspiration from qualitative research. By taking a more socio-constructivist approach, concepts such as credibility, dependability and confirmability have been introduced to WBA. Govaerts *et al.* (2007) consider that 'multiple interpretations' of a student's performance may be equally valid and put together 'present a rich and detailed report of competencies and the situation-specific behaviours' (p.254). For Gingerich, Regehr and Eva (2011) adequate sampling might turn the bias of raters into idiosyncrasies providing a more rounded perspective of a student. This presents a different way WBA can be viewed loosening the ties to objectivity and a psychometric discourse and resonating more with authenticity and holism. Schuwirth and van der Vleuten (2006) recommend a shift of focus from assessment as a 'measurement of the outcome of an educational activity to assessment as an integral part of education' (p.299). This line of inquiry will be followed up further.

2.3.3 The single true score

The psychometric discourse rests on a concept that students possess abstract, stable and latent dispositions that can be measured. There is, however, a large body of literature that challenges the possibility of producing a 'single true score' about a student's ability on a given competency statement (be it knowledge or techniques or professionalism) and maintain validity and reliability (Durning *et al.*, 2013; Ginsburg *et al.*, 2010; Gruppen *et al.*, 2018; Hodges, 2013; Schuwirth and van der Vleuten, 2012).

If the primary aim of grades is to distinguish one student from another then it is critical to know whether competence in a student is a stable and generic trait of that individual across contexts, or a state that varies with the situation and thus whether any variance of grades is 'noise' or 'signal'[7] (Durning *et al.*, 2013). Following from the assumptions described earlier within the psychometric discourse, Govaerts and van der Vleuten (2013, p. 1165) consider further issues. One is that 'competence, as inferred from performance, is a fixed, permanent and de-contextualized attribute' and that a student's performance is stable across cases and situations. The other is that 'performance can be 'objectified' and assessors, if they were only capable to do so, would be able to rate and observe some true level of performance'. In health care, however, they argue, learning is a non-linear, non-deterministic and dynamic process. For Schuwirth and van der Vleuten (2012) evaluation of performance is an attempt to measure and predict the interaction of the student and the environment not of a stable characteristic and therefore interaction effects should be seen as meaningful signals not 'noise'. Hence, in support of Gingerich, Regehr and Eva (2011), the different results of assessors can be a valid and reliable measure of a student's performance.

Whitehead *et al.* (2015) pick up the paradox contained in the challenge to medical education by society to find ways to assess the qualities of doctors such as professionalism and empathy. These attributes are the very things that are elusive to measurement and yet the pressure for accountability has led to the dominance of psychometric assessment systems. While the ICBME consortium acknowledge the problem of context specificity and the inherent uncertainty and imprecision of WBA in coming to numerical decisions, Gruppen *et al.* (2018) propose to remedy this through the systematization of subjective judgements such as:

> behavioural anchors for rating scales, detailed descriptions of what 'competent' performance looks like, and faculty development to calibrate faculty to a common set of criteria. (p. S19)

---

[7] 'Noise' denotes an error in the measurement system and 'signal' indicates true variance in the student's performance.

The calibration of not just the assessment instrument but also the assessor, are both critical in this approach. This leads me to turn my inquiry to arguments on the role of the assessor.

2.4 The Assessor

In agreement with Gruppen *et al.* (2018) above, Lockyer *et al.* (2017), Shumway and Harden (2003) and Williams, Klamen and McGaghie (2003) all consider that the onus of the validity of the assessment process lies with the individuals conducting the assessments. If the assessment instrument is only as good as the person using it then effort would need to be spent on reducing assessor errors and biases. In this section I will consider the literature on assessor bias and the role of training and then expand into the place of assessors' judgement and interpretation and recommendations that WBA must support learning as well as grading.

2.4.1 Causes of Rater bias

A number of factors have been identified that impact on why assessors may not produce reliable scores and why inter-rater reliability might be compromised. It might be down to the raters' own competence and skills; variable frames of reference; a local culture; anticipation of possible reactions by students to marks given; and cognitive factors such as subconscious stereotypes or personal idiosyncrasies (de Jonge *et al.*, 2017; Holmboe, 2015; Kogan *et al.*, 2011). Biases such as halo effects or leniency are also factors reported (Gingerich, Regehr and Eva, 2011; Pangaro and ten Cate, 2013). Another view introduced by Shay (2004) and Harman and McDowell (2011) is that assessors see their role as encouraging development rather than producing a discriminatory grade. In relation to the health professions, Yorke (2005) proposes that supervisors tend to be generous in their assessment as it aligns with their sense of nurturing. This is not bias but an alignment of the assessment construct with the values of the assessor.

The issues raised in the previous sections also challenge perspectives on assessor bias in several ways. The first is that competency statements are

explicit and universal and so understood equally by assessors. The second is the concept of there being a single 'true' score and that the multiple perspectives of different assessors are noise (or bias) and need to be reduced. The third is the quest for objectivity in the form of numbers and that this equates to reliability and validity. Fourthly, the association between objective assessment and standardized instruments is doubtful (Hodges, 2013).

2.4.2 Assessor Training

Whether the training of assessors is there to calibrate assessors and thus improve inter-rater reliability by eliminating personal biases or is there to help assessors, and groups of assessors, understand their personal biases through developing shared mental models, is debated. As discussed earlier, the idiosyncrasies of the assessors can be viewed as valid interpretations as opposed to biases.

Kogan and Holmboe (2018) consider that the training of assessors is necessary and collates evidence that various programmes such as 'behavioural observation training', 'performance dimension training' and 'frame of reference training' have all had a positive impact on reducing assessor bias. The research of Hodwitz (2018) showed training improved assessors' understanding of the role of judgement, improved assessors' confidence in scoring and utilizing assessment tools, but it did not affect inter-rater reliability. A randomized controlled trial carried out by Cook *et al.* (2009) on the impact of assessor training on the use of a particular assessment instrument showed no improvement in inter-rater reliability. The conclusions of Gingerich, Regehr and Eva (2011) and de Jonge *et al.* (2017), were that any solutions attempting to bolster assessor objectivity have had little impact and that the objective evaluation of performance in complex settings is not possible and perhaps not even desirable. If, as Lockyer *et al.* (2017) propose, validity is not an inherent property of an assessment instrument but is in the hands of the assessor then, as in any other skill, I propose, it is the process of assessment that must be learned, not the calibration of the assessor to the instrument.

When it comes to marking, Shay (2005) proposes that if internally developed,

marking criteria do reproduce the norms and rules, the tacit standards of the group designing them. Meddings (2017) also describes how over time assessors learn to mark by marking and how communities of practice support a cohesive approach to marking. Crossley (2013) suggests that while calibration training of raters did not reduce variability in their interpretation of performance, discussion around performance and aligning assessment instruments with existing value based cognitive frameworks did lead to more shared interpretations. This could explain the positive outcomes of assessor training that Holmboe (2015) reported.

Govaerts *et al.* (2007) and Boudreau, Cassell and Fuks (2018) argue for training for teachers to go beyond the effective use of the assessment instrument to the development of appropriate, open, supportive and non-defensive relationships with students. As Bleakley (2015, 2019) and Berendonk (2013) point out the high behavioural and cognitive demand of active performance is a vulnerable position that creates emotional tension for students and for teachers and assessors, changing with each patient and situation.

Van der Vleuten's (1996) position, as explained already, is that subjective, global rating scales based on experts' judgements can provide reliable scores and might be better able to discriminate between levels of expertise in students. It follows that the concept of assessor bias might need to be reframed and a closer look taken of the place of interpretation and judgement in WBA.

2.4.3 Interpretation and Judgement

Pre-specified competency statements rely on precision in the language used to provide commonality amongst assessors, with Jessup (1991, p. 134) going so far as to state that 'the overall [CBME] model stands or falls on how effectively we can state competence and attainment'. The assumption, however, in standardized outcomes is that the language in which they are written is transparent (Crossley *et al.*, 2011; Morris, 2019) and free from mis-interpretation or ambiguity for those applying them (Dunne, 2005; Grealish, 2015). Hodge, Mavin and Kearns (2019) critique the very process of interpretation that occurs as competence is translated from the analysis of the

practice by the profession through to competency statements held within text, and then into curricular design, teaching activity and assessment instruments. They question the authority that is ultimately invested in the texts produced. Any valid measurement must rely on the assessor's interpretation of the competency statements (Govaerts *et al.*, 2007; Yorke, 2011a) returning us to question how standardized they can be.

Sadler (2009) supported by Kilgour *et al.* (2013) report that word-based outcomes and criteria 'may be interpreted differently by different teachers' and that 'they can also be interpreted differently by the same teacher in different assessment contexts' (p.169). Further, assessors attach importance to different aspects of students work (Bloxham *et al.*, 2015; O'Hagan and Wigglesworth, 2014). While criteria might be presented as discrete entities and thus assumed to be conceptually distinct they will be merged differently by teachers as their boundaries are actually 'fuzzy and situationally dependent' (Sadler, 2009, p. 69). Ginsburg *et al.* (2015) consider this might lead to construct 'mis-alignment' constraining authentic depictions of students' performances and put a strain on the transparency of the assessment process. The standards an assessor applies to a student's work are often emergent, dependent on the context, based on personal values and are often tacit (Giloi, 2014; Govaerts and van der Vleuten, 2013; Orr and Bloxham, 2013).

In arguments around assessor bias Hager and Butler (1996) insist that subjectivity is not the same thing as bias.

> Objectivity is the intelligent learned use of subjectivity, not a denial of it. In the judgemental model of assessment it is the assessor who delivers objectivity, not the data. (p. 372)

They describe assessment as the application of an assessor's judgements based on learned personal standards. This implies that objectivity will be better achieved by trusting more in judgement in context rather than measurement as judgement as it will be the case that non-variable instruments are unlikely to capture the normative judgements being made by assessors.

An outcome of the analysis by Gingerich, Regehr and Eva (2011) was that assessor error may be due to the competency frameworks against which they

are asked to make judgements; they may be 'incongruent with the cognitive processes used by humans to perform judgements' (p. 55). Assessors, they say, may be making consistent yet different categorizations of students' performance. Not just the language but the representation of atomized competencies in numeric scales is what Regehr *et al.* (2007) concluded was being resisted by assessors. Govaerts *et al.* (2007, p. 252) propose that assessors are not 'passive measurement instruments' applying strict performance criteria laid out in assessment instruments, but active processors of information applying personal meaning, perspectives and values to the assessment. For Broad (2000), competency statements might not assist assessors as they do not make apparent the value system that underlies them. As discussed earlier, the idiosyncrasies of assessors as they make sense of complex clinical scenarios are said by Gingerich *et al.* (2014, p. 1055) to be 'legitimate experience-based interpretations'. If assessors' differences of opinion are 'signal', then the aim of assessing performance is not to come to a measurement but to leverage the assessors' reasoning and decision making process.

Sadler (2009) sums up the tensions that can ensue in WBA due to the complexity of the judgements needed in multi-criterion assessments, returning us to the tension between atomization and holism. He proposes that judgements are not always made at the level of individual criteria; criteria do not hold inherent meaning but are interpreted; and assessors are aware of discrepancies between global and analytic conclusions but they can remain tacit. The implications of these arguments are that assessors are making a global judgement of a student's competence. Dunne (2005, p. 375) considers that competency statements '[dis-embed] the knowledge implicit in the skillful performance of the characteristic tasks of the practice from the immediacy and idiosyncrasy of the particular situations in which it is deployed'. These would need to be re-embedded by the assessors as they make inferences about competence from multiple observations of the student's performance (Grealish, 2015; Hager, 2017). The research of Bloxham, Boyd and Orr (2011) found evidence showing that assessors come to a decision about a general grade and use the published criteria to check. They state:

> markers are not cynically referring to criteria post hoc in order to defend judgements, but are using them to refine 'hunch' decisions. (p.662)

Berendonk (2013) also reports assessors using hunches or 'gut feelings' as they apply their own professional and personal knowledge of the discipline that takes into account the student and patient in context. Shay (2004) supports the argument that the professional judgement assessors are using is not bias but a form of rationality. Not objectivist or relativist, but contextual and experienced and value based. Rather than standardizing and atomizing, rather than 'tightening up', Eisner (2004) calls for a qualitative intelligence to be developed including a refining of the assessors sensibilities, of perception, attentive to relationships, a flexible purposiveness open to uncertainty.

The lack of evidence of objectivity in the measurement of grading complex assessments such as in WBA, leads us to conclude that it is at root a socially constructed activity involving a process of interpretation and judgment. Validity is therefore not a quantitative procedure but an interpretive process involving the complex application of judgement influenced by the interactions between assessor and student and the context in which the assessment occurs. Rather than inter-rater un-reliability being a problem of implementation or interpretation (Boyd *et al.*, 2018) as already discussed, an assessor's judgement can be viewed as a valid interpretation of a student's competence.

Rotthoff (2018) points out, however, that to reach a valid judgement and transform the assessors observations into a grade, given that is a requirement of the quality assurance discourse, does need some degree of familiarity with the defined constructs and dimensions of the assessment process and framework along with a reflected discussion with other stakeholders in order to come to shared mental and value-based understandings. Otherwise, he proposes, there is a tendency to de-value the assessment instrument and give above average ratings. The aim of assessment therefore will be lost for all stakeholders.

2.4.4 Assessment of and for Learning

> Edumetric: used of approaches to assessment which focus on authentic tasks and acknowledge cognitive complexity. Unlike psychometric

approaches, which emphasize differences between learners on the normal curve, edumetric approaches focus on individual learning or within-individual growth (Gillies, 2010).

This definition of 'edumetrics' clarifies one of the issues of what it is we may need to measure in WBA. A psychometric approach is particularly concerned with differentiating between students, an important concern in 'high stakes' assessments that will determine entry to higher-level training or employment. Edumetrics is more concerned with the development of the individual student. This is a further example of the instrumental discourse as assessment-of-learning coming into tension with the socio-constructivist discourse as assessment-for-learning.

That assessment must support learning, self-directed learning and learning for life before issues of reliability or validity especially in longitudinal assessments is proposed by many educationalists. Boud (2009) and Higgs, Loftus and Trede (2010) argue that good assessments can be used to determine grades, but it is crucial that assessment fosters and helps rather than hinders the development of the required attitudes and dispositions students will need in their professional life. Schuwirth and Van der Vleuten (2011a) urge that the primary requirement of observation of performance should be developmental with the observer accompanying the student on their learning journey and summative decisions should be secondary. Picked up by Dornan (2012) and Winstone and Boud (2020) is the concern that assessment-of-learning can consume assessment-for-learning as the application of judgement on the part of the assessor as to the student's competence is overwhelmed by metrics and standardization.

Summary

This literature review of the multiple discourses influencing WBA has been important in setting the context and relevance of this study. The key issues that have emerged from this chapter so far relate to: (a) the tensions between standardization and authenticity and between atomization and holism and the different ways in which WBA will be enabled or constrained by these constructs; (b) whether the assessment instruments of WBA can lay claim to validity and reliability; (c) the factors determining assessors' interpretations of competencies

and the judgements they are making and (d) how assessment-of-learning and assessment-for-learning can be in tension.

Trede and Smith (2012, p. 193) suggest that 'the means to enhance WBA is not to promote a dominant discourse but rather to assist assessors (and I add all stakeholders) to reconcile and work within these tensions'. However, the CBME construct remains rooted in discourses of product, performance and psychometrics (Hodges, 2012; Landri, 2012), of de-contextualised individual competencies (Fenwick and Nerland, 2014), of context being limited to settings or places (Bates and Ellaway, 2016; Evans, Guile and Harris, 2011) and focuses on the individual student (Bleakley, 2006). Sebok‑Syer *et al.* (2018) question the very fundamental assumption held within CBME that places the student as the focus of assessment attention reporting instead on the interdependence of the assessor and assessee.

Perhaps, I conclude, it is the very construct of CBME that needs to be debated further. Boyd *et al.* (2018) challenge those supporting CBME as promoting a discourse of 'infallibility'. Their analysis of the literature found that any critical voices of CBME were being silenced and conceptual concerns around assessment were being 'reframed as two practical problems: implementation and interpretation' (p.45). Docherty (2017) also questions the epistemology of CBME as being insufficient to explain and develop the process of assessing competence to practise. According to Boud *et al.* (2018, p. 1107) 'assessment as a field of investigation has been influenced by a limited number of perspectives' with the predominant focus being the assessment instruments and attempts to improve their reliability and validity. To get to a deeper level of understanding of WBA I needed to de-centre measurement (ibid) and the tasks students perform as described in the learning outcomes. I needed to theorize WBA further, and consider the nature and entwinement and whole phenomenon of assessment as the unit of analysis (Fenwick, 2012). In the next section I present 'practice theory' as that different theory with which to view WBA. Sebok‑Syer *et al.* (2018) and Boud *et al.* (2018) report that a practice theory approach has rarely been used in research into assessment especially within medical education.

2.5 The Practice Turn

In this section, I present Theodore Schatzki's approach to 'practice' as described in 'The Practice Turn in Contemporary Theory' (2001) and developed in Schatzki (2012). I also draw on Kemmis and Grootenboer (2008) who have developed the theory of 'practice architectures' in order to provide a more empirically manageable approach to understanding practices. My aim is to view WBA as a 'practice' in its own right. Seeing WBA as a socio-cultural-material, embodied and interactive phenomenon that gives space for the human voice and human values as well as non-human actors (e.g. artefacts, buildings) in time and space opens up a different and wider perspective from that offered by CBME.

2.5.1 Schatzki's Definition of a Practice

Schatzki (2002) describes practices as:

> … organised nexuses of actions. This means that the doings and sayings composing them hang together. More specifically, the doings and sayings that compose a given practice are linked through (1) practical understandings, (2) rules, (3) a teleo-affective structure, and (4) general understandings. Together, the understandings, rules and teleo-affective structures that link the doings and sayings of a practice form its organisation (p. 77).

Before I de-construct and expand on this definition I must add that being a socio-cultural-material phenomenon 'practice' is a contested term embedded in tradition, diverse and multi-dimensional, interpreted differently by different stakeholders and understood differently depending on the different intellectual and scholarly traditions of the theorists. I do not have the space to present the range of ideas contained in the literature but refer the reader to Schatzki (2012) and Mahon *et al.* (2017) who map out the history and theoretical terrain of practice theory. I have extracted from the literature what I have found most relevant to help expand our understanding of WBA. To do this I will deconstruct Schatzki's definition of practice as quoted above.

2.5.2 Sayings, Doings and Relatings

According to Schatzki a practice is an:

> 'open-ended, spatially-temporally dispersed nexus of doings and sayings' (2012, p. 14)

By open-ended is meant any number of activities, or paths of activities, that can take place somewhere in some space at some point, held together (a nexus) by 'sayings, doings and relatings'. It is Kemmis and Grootenboer (2008) who have added the expression 'relatings'. Practices are always 'located in particular sites and particular times' (Kemmis *et al.*, 2014, p. 33) with meaning established through the network of interwoven activities. It is that the sayings, doings and relatings 'hang together', happen together and are connected in some kind of coherent way that give a practice its distinctiveness.

That a practice is constituted by 'sayings' refers to what people think and say in words spoken and written. A practice is not just represented but also shaped by discourses, by forms of thought and language and beliefs and values including cultural and traditional distinctions and specialist discourses that make it understandable and interpretable as a specific practice (Kemmis, 2009). Boud (2000) and Price *et al.* (2011) highlight assessment in particular as carrying out this role within education. It must be noted that discursive and cultural arrangements can enable and also constrain particular sayings as noted earlier in the work of Ho *et al.* (2011, 2012, 2014).

Gherardi (2019) presents a view that removes us from the more cognitive approach describing practices as 'performed through a 'sapient' body that knows through the senses, and that accumulates in the body a capacity to act in the world' (p.151). The body, for Schatzki (2001), is the meeting point of mind and activity hence the emphasis in practice theory on embodiment. The role of the body and the senses in the way we make meaning and mediate the world has been well developed by Merleau-Ponty (2013). He proposes that

> the visible is what we grasp with our eyes; the sensible is what we grasp through our senses (p.7).

This quote links back to the role of judgement in WBA and that the personal

meaning, perspectives and values assessors are applying to students' work are often tacit while Landri (2012) and Schatzki (2017) describe practitioners (as assessors) as needing to develop an aesthetic understanding.

That a practice is constituted of the 'doings' of individuals refers to both a physical and material space in which activities take place but also within an economic context of resources, production and exchange. Material and economic arrangements and responsibilities shape the doings of practice by affecting what, how, when, or by whom something can be done (Mahon *et al.*, 2017). Gherardi (2009b) sees tools and artefacts and objects as having agency by embodying knowledge and anchoring practices. For Fenwick (2012), seeing the materiality of a practice as not separate or distinct from the human designers or users helps to lessen the privileging of the human actors. As example, I suggest, assessment instruments as artefacts and their psychometrically determined outcomes, as described in the previous chapter, carry a privileged position.

A practice also consists of the relationships that occur between people, relations of belonging or not, of inclusion or exclusion, between different roles formal or chosen (Kemmis, 2009). As Gherardi (2009a) proposes, a practice is not only formed by what people do but the ways they do things together, going beyond the knowledge, experience, intentions and actions of each person as an individual agent (Kemmis *et al.*, 2012). A practice consists of the forms of knowing generated through the relationships among the participants and also material arrangements as well as the everyday routines and interactions. Social and political arrangements or resources shape how people relate in a practice, not just to each other but also to non-human artefacts.

Reich and Hager (2014) suggest that practices are not stable or homogeneous but exist and evolve in historical as well as social contexts and are shaped by complex social and political forces. Further, the traditions of a practice, Mahon *et al.* (2017) describe as carrying the imprints of prior sayings, doings and relatings and so too can new sayings, doings and relatings transform practices. It is the characteristic arrangements of discourses, activities and social relationships that make a practice distinctive.

2.5.3 Rules, Understandings and Teleo-affective Structures

Returning to Schatzki's definition of practice, doings, sayings and relatings only belong to a given practice when they convey the rules, the principles, directives or instructions that direct people to perform or not perform an action, or to act or not act as expected. Rules can be both explicit and implicit. Fenwick (2012) explains that there are codified ways of working as well as ways of working that can be seen to bend the rules but which make the codified practice work. At the same time some rules are more implicit or tacit as in the way information is coded or interpreted, the ways instruments are used or conversations held. Eraut (2000) proposes that tacit rules often underpin formal decision-making.

For Schatzki, understandings can be both practical and general. Practical understandings pertain to the specific actions that compose the practice. They combine not just knowing how to perform an action but also the sense of the action, how to recognise the action and how to respond to it. General understandings refer to beliefs and concerns that 'tint' practices (Lammi, 2018) and can include abstract senses such as the value or worth of something or a shared understanding of the significance of work being done. Understandings are infused and are expressed in people's sayings, doings and relatings.

By teleo-affective structures Schatzki is referring to the project of a practice, the normatized or ordered ends of a practice. He describes sayings, doings and relatings as organized through purposeful intentions (teleo) and enjoins the emotions and moods expressed by the people in the practice. Teleo-affective structures embrace the investment and motivation of the people in the practice, are appropriate to a specific practice and govern what is specified by rules or understandings (Schatzki, 1997, p. 3). But practices are also emergent, dynamic and entangled, changing and evolving in ways that cannot be specified in advance (Hager, Lee and Reich, 2012). Thus a practice is organized not just by purposeful structures but also happen-chance structures (Schatzki, 2012). Sayings, doings and relatings can be seen to pertain to a practice if organized and also linked by rules, understandings and normativized or acceptable ends and affectivity.

Summary

Returning to the beginning of this thesis, Eva and Hodges (2012, p. 914) drew our attention to the two different epistemological approaches to assessment that are most commonly referred to in the literature. From one angle is the instrumental-technical approach demanding objectivity in the form of measurement and from the other a socio-constructive approach that recognizes context and the subjective judgement of the assessors. However, as I have described, when assessment is viewed as a practice, cultural-discursive, material-economic and social-political arrangements, circumstances and conditions become foregrounded. It is the characteristic arrangements of discourses, activities and social relationships that make a practice distinctive.

For Kemmis *et al.* (2012) using their concept of practice architectures, a practice is constituted in a semantic space, a material space and a social space. For example, Boyd and Bloxham (2014) describe grading in assessment as including both the semantic space of pedagogy and the subject discipline along with a practical wisdom that is formed in the social workplace of the assessor. This practice lens gave me the possibility of viewing assessment-as-practiced (Boud *et al.*, 2018), illuminating the different relationships and the different mediating conditions and entwinement of the sayings, doings and relatings of the different stakeholders.

It is the combined theories, the instrumental-technical, socio constructive and also practice that form the framework for my exploration of WBA in institutions teaching Chinese Medicine. My interest is in how WBA is being enacted and what might be 'enabling or constraining' (Kemmis and Grootenboer, 2008) it as a practice. In the next chapter I will lay out my methodological rationale and design that takes into account how I have created access to the manifestation of the theories described above.

Chapter Three: Research Design and Methodology

Introduction

As illustrated in my professional context, my own perspectives on assessment
have emerged from my experiences of being an accreditor, manager and
clinical supervisor. These multiple roles provided an insider and outsider
perspective in relation to the different stakeholders I was researching and
presented both opportunities and challenges. This chapter will show how this
tension was managed through the design and implementation of this study
including selection, recruitment and choice of location for interviews and into the
various stages of my analytical process as I constructed meaning. The ethical
considerations that permeated the whole process will be clarified including my
reflections especially in relation to interviewing.

3.2 Research Design Rationale

In order to explore my research questions, I chose a qualitative method that
could see beyond the intrinsic qualities of assessment instruments (de Jonge *et
al.*, 2017) to the perspectives of the multiple stakeholders involved in
assessment. This involved taking into account the accreditation documents that
set out standards for institutions in each of the three countries, the decisions of
the institutions running the courses, and the experiences of the supervisors
carrying out the assessments. Documents to consider included accreditation
standards and competencies, institutional course and clinic handbooks and the
assessment instruments used by the assessors. In order to understand the
socio-cognitive filters or frames (Goffman, 1986) that the people acting on the
assessments might be applying I interviewed both the managers who had
written the assessment documents or were overseeing the process and the
supervisors who carried out the assessment. The setting in which the WBA took
place was also a factor to consider and was brought about by my physically
viewing the institutions and teaching clinics. A qualitative research method
framed as a multiple case-study (Stake, 2005; Yin, 2014) and employing a
socio-interpretivist approach I considered as the most appropriate to guide and

contain the study. My research data has come from three institutions in different countries as outlined below in Table 1.

3.3 Case Study

Being both exploratory and explanatory, the case-study approach seeks for an in depth and rich description of a phenomenon in its real-life context in order to shed empirical light on a theoretical concept that is perhaps more complex than it is sometimes presented. The multiple embedded case study design described by Yin (2014) allowed for the use of multiple sources of evidence (see Table 1). The documents that I sampled were not there to corroborate findings but were a key source of evidence towards developing my findings. As Prior (2003) proposes documents are not inert but manifest a social and political context and are imbued with the institutional and personal understandings, beliefs and affiliations of the authors. Critically, they also gave access to a different group of participants not able to be interviewed.

The unit of analysis was the assessment processes in three institutions within the context of the accreditation and curriculum documents and as described by the managers and supervisors.

Table 1 Multiple case design

	Case 1	Case 2	Case 3
Case A	Accreditation documents in USA	Accreditation documents in Australia	Accreditation documents in UK
Case B	Institutional documents	Institutional documents	Institutional documents
Case C	Views of Managers	Views of Managers	Views of Managers
Case D	Views of Supervisors	Views of Supervisors	Views of Supervisors
Case E	Clinic architecture	Clinic architecture	Clinic architecture

By choosing a number of institutions (cases 1-3) and multiple sources of evidence (cases A-E) my aim was to apply a systematic examination of similarities and differences within and across cases (Brinkmann and Kvale, 2015) allowing me to examine how a phenomenon might behave differently in different contexts.

To understand the issues of case study research in terms of its generalizability and transferability I turned to Stake's (1995) description of an 'instrumental' case study in which the issues dominate as opposed to each case being individually critical. My aim was to use the common and variable data from the three institutions to provide critical analytical insight into assessment processes. These insights and the concepts and theories generated through the research (remembering that they emerge from my experience and perspective as the researcher (Cheek *et al.*, 2018) and my immersion in and engagement with the data (Polit and Beck, 2010)) could then be transferable to other CM institutions and to other assessment situations (Robson, 2011; Silverman, 2017). Flyvbjerg (2006) considers that:

> The goal is not to make the case study be all things to all people. The goal is to allow the study to be different things to different people (p. 238)

In other words any generalisation is thus subjective on the part of the reader (Stake, 2005). Polit and Beck (2010) propose that any generalisation from case study research is best considered as a working hypothesis rather than a conclusion.

3.4 Data Collection

In this section I will describe my choice of countries and recruitment of institutions along with the rationale for selection and the problems encountered. The process of interviewing, the questions asked and details of conducting the interviews are described concluding with a reflection on the role of the 'insider' researcher

3.4.1 Inclusion Criteria

The first objective for my research was to consider what institutions I would need to gain data from. My experience as an external examiner of a number of institutions in the UK had led me to see that, because of accreditation and validation requirements, WBA was being applied similarly. I decided, therefore, to include the UK but also to expand my investigation to other countries that had different histories and regulations and processes. If the meanings and interpretations individuals make are embedded in cultural and social networks of language, artefacts and symbols (Bleakley, Bligh and Browne, 2011), my aim of looking at different countries, different landscapes (Denzin and Lincoln, 2008) would provide, potentially, a diversity of practices and experiences in thinking about WBA. The following shows my inclusion criteria and rationale.

Table 2 Inclusion Criteria

Organisation	Criteria	Rationale
Country	English speaking	my main language is English and I needed to understand the documents and conduct in-depth interviews
	CBME a main aspect of medical education	this would have influenced the accreditation processes
Institution	Professional entry level training in CM	as opposed to medical acupuncture as students would have already completed a biomedical training
	Course was a full-time 3-4 year CM training	this was to allow for cross comparison
	Institution provided the final clinical training	as opposed to external placements to allow for viewing of the locus of assessment
	Validated within HE systems	have had to manage issues of standardization from an educational perspective
	Accredited by relevant professional organisation	have had to manage issues of standardization from a professional perspective

The countries that met these criteria were the USA, Canada, Australia and the UK. I excluded Canada as I had no contacts there. Of the institutions, three in Australia met the criteria and four in the UK. Of those in the UK two were eliminated as a result of the issues of being an 'insider' researcher. For one I had an official relationship as an external examiner and the other I felt that competition with my own course might not be conducive to the open style of

interviewing needed including, on my part, anticipating their response. In the USA there were many courses that met the inclusion criteria.

3.4.2 Recruitment Process

Knowing personally and professionally a number of course managers in Australia and the UK, initial enquiries went to two institutions in each country and responses were positive enough for me to proceed with developing the research proposal and applying for ethical approval.

I sent the developed and official request to the course managers of two institutions in the UK and three in Australia (Appendix 1). In the USA, where there were many courses I might have researched, my contacts were more tenuous and colleagues forwarded the official request. As I also required the course managers to disseminate my request to the supervisors they were also sent the Teacher Participation Letter (Appendix 2).

On following up these formal letters access was not as easily granted as I had anticipated so my choice of institutions was limited to those who gave agreement, one in Australia and one in the UK. In the USA it was a matter of timely response to my colleagues' enquiries and one institution that met the criteria came forward. The reasons why some institutions were not as cooperative as I had anticipated were varied. One course leader whose original contact was positive was happy to cooperate individually but did not feel they had the authority to disclose curriculum documents developed in conjunction with their affiliated institution in China. The documents, however, were an essential part of the research design. Another course leader gave a definite rejection citing that they were not interested in this sort of research. One institution had decided to close the CM course and the course leader no longer wished to participate.

In considering whether to proceed with one institution in each country I turned to Yin's (2014) recommendations. In qualitative research it is not the numbers of different cases that matter as much as the depth of information that is revealed. Breadth and depth would be gained from considering three

stakeholders, multiple documentation and also multiple interviews within each institution. As Yin recommends, the number of cases or 'replications' needed is also more to do with the degree of certainty the research requires and as this field has not been researched before in CM, the three cases could provide the commonality and variability needed to answer my research questions (Flyvbjerg, 2011; Silverman, 2017). There was a common history of the three institutions that I felt conducive to providing the similarities to bind them as multiple cases. They had moved from being independent schools owned and run by groups of practitioners into universities or HE settings, with the consequential uptake of new management structures and input of different accreditation and validation requirements. At a management level the three institutions were also quite different with different corporate structures impacting on their assessment processes. Understanding the country-specific and institution-specific contexts due to the different stakeholders, their history, social structure and relationships that influenced their curricula would provide the basis for comparison.

3.4.3 Data Sources

The gathering and analysis of accreditation documents from each country (Table 3) was needed to understand how the regulatory context of CM within each country (see below) and the ethos of assessment was impacting on WBA, what was required of institutions in relation to WBA and what role the construct of CBME played in these requirements. Gaining access to the course, clinic and assessment documents from the course managers of the institutions in each country (Table 4) was needed to understand the ethos as well as the published position of assessment within the institution and what role the construct of CBME played. Having agreed a time to attend the institutions, I interviewed the managers and supervisors who consented to participate and observed the clinic environments. While the documents provided the official position on WBA, the interviews with both managers and supervisors allowed access to the life-world of those involved in WBA.

National Regulatory Context of CM in each country.

The majority of states in the USA require practitioners of CM to have completed an accredited programme and to hold the National Certification Commission for Acupuncture and Oriental Medicine (NCCAOM) examination or certification. The Accreditation Commission for Acupuncture and Oriental Medicine (ACAOM) accredits CM programmes. ACAOM is recognized by the United States Department of Education (USDE). NCCAOM certification is contingent on completion of an approved educational program, passing the NCCAOM certification examination and completing the Clean Needle Technique course. There are also regional accrediting bodies (not related specifically to CM), which accredit institutions as opposed to programmes, although this is not a legal requirement. ACAOM may also provide that institutional accreditation.

In Australia, the practice of CM is statutory controlled by the Australian Health Practitioner Regulation Agency (AHPRA) and title is limited to state-registered practitioners. Practitioners must be registered with the Chinese Medicine Board of Australia (CMBA) and the CMBA Accreditation Committee, which carries out accreditation of programmes, was established by the Health Practitioner Regulation National Law. Institutions teaching Chinese Medicine are also governed by requirements of the Tertiary Education Quality and Standards Agency (TEQSA). It is the institution that gains accreditation and students gain entry to the profession if they pass their programme.

In the UK, there is no statutory regulation for CM and practitioners operate under 'Common Law'. Most non-medical practitioners belong to a voluntary, self-regulated professional body such as the British Acupuncture Council (BAcC). The British Acupuncture Accreditation Board (BAAB) is a not-for-profit company carrying out accreditation of acupuncture programmes or institutions affiliated to the BAcC. Accreditation is a voluntary process for institutions in the UK. In 2010, while not legally required, the BAAB brought its Standards of Education and Training for Acupuncture Programmes (SETAPs) in line with requirements of the Health and Care Professions Council (HCPC) to provide a level of ratification of its own processes. Academic validation with a University is

also a route that can be taken by an institution but is not an accreditation, legal or educational requirement unless the institution wishes to award a degree. If so they will need to meet the standards of the Higher Education Funding Council of England (HEFCE). In the UK it is the programme or institution that gains accreditation and students gain entry to the professional body if they pass their programme.

Table 3 Accreditation Documents Used in the Study

	USA	Australia	UK
	ACAOM Standards and Criteria Manual 2016 (2019) (ACAOM, 2016) (ACAOM, 2019)	CMBA Chinese Medicine Accreditation Standards (2013) revised Dec 2016 (CMBA, 2016; CMBA, 2019)	BAAB Accreditation Handbook (2010) revised 2016 (BAAB, 2016)

Table 4 Institutional documents Used in the Study

Institutions	USA	Australia	UK
Possible institutions (after criteria applied)	numerous	3	2
Number researched	1	1	1
Institution documents	Course and Clinic Handbook	Course and Clinic Handbook	Course and Clinic Handbook
Assessments	Clinical assessment instrument	Clinical assessment instrument	Clinical assessment instrument

Interviews with Managers and Supervisors took place during 2016-17. At each institution I made myself available over a 4-7 day period.

Management. I interviewed 3 staff members in each institution differently responsible for the development, design or management of the course documents, clinical teaching or the assessment arrangements. Interviews with each manager took between 40 and 50 minutes. In two institutions the clinic director as manager was also a supervisor and their input has been included in both data sets. In one institution a meeting was conducted by Skype due to a lack of availability.

Supervisors. I relied on the course leader or clinic manager to provide an introduction to the supervisors. The only inclusion criteria for supervisors was that they needed to be those who conducted the WBA and willing to be interviewed on the days I was available. 14 supervisors volunteered with each interview taking from 30-45 minutes except for one of 9 minutes, which will be discussed in the section on ethics. One meeting was conducted by Skype due to a scheduling issue.

Location. As it was the life-world of the participants that I wanted to gain access to, interviewing in their place of work was part of the research design (Dall'Alba and Sandberg, 2014). Interviews were conducted in unoccupied clinic or discussion rooms on the days the supervisors were in clinic and took place in the times the supervisors had free before or after or in the middle of their work. As the real world impinged on the data gathering it provided insight into the nature of the clinic day and the supervisors personal relationships with students and with other faculty members. On occasions, the limited time available for the interview and the occasional interruptions did impact on the information being received curtailing more depth to some aspects of discussion. However, this was balanced by the perceptions gained from being immersed in the supervisor's world.

The common features and differences afforded by the three countries, three institutions, 9 managers and 14 supervisors produced the breadth and depth needed to meet my research aims. Below is a summary of the demographics of the 14 supervisors interviewed (Table 5).

Table 5 Supervisors' Demographics

No of interviews	13 interviews (only the demographics of the 9 minute interview were included)
Timing	30 – 45 minutes
Years since graduated	5-38 years

Trained in own institution	7 (50%)
Years doing clinical supervision	2-22 years
Formal teacher training	2 had some formal teacher training 1 had some teacher training in another discipline

3.4.4 Interviewing

The approach to interviewing that I adopted was based on that described by Brinkmann and Kvale (2015) with the aim to gain insight into the participants' roles, experience, views, opinions and feelings. I chose a semi-structured, focused and face-to-face interview method.

My theoretical concepts formed the basis of my interview questions (Appendix 3). The application of similar questions to each participant in each group allowed for a level of comparison and contrast as recommended by Brinkmann and Kvale (2015). The questions to the course managers were piloted on a member of staff at my own university who had been involved in our curriculum development and the supervisor questions were piloted on a member of staff who was a clinical supervisor.

3.4.5 Interview questions

With the managers my aim was to understand the structure and system around clinical assessments and their personal involvement in their development.

> This study hopes to explore the way that Chinese medicine (CM) Higher Education institutions in the West assess the professional practice of their students

Table 6 Interview Topics for Managers

1. General context of the course – history and organisation of the institute
2. What do the managers understand by the stake holders (accreditation and validation) impact on assessment choices
3. What is the relationship between the management and the supervisors – communication and decision making
4. How do the managers understand the issues behind the application of assessment instruments to practice-based assessments – learning outcomes, standardization
5. Using the idea of 'failing' how are the criteria administered

The questions to the clinical supervisors were summarised in the information
sheet with an open statement:

> I wish to understand your personal experience of clinical supervision and
> also assessing students practice.

Table 7 Interview Topics for Supervisors

1. General context of the supervisor
2. How do they understand their development of their role
3. How do they understand assessment
4. What do they understand about learning outcomes
5. What is their experience of using the assessment instrument
6. What community of practice do they identify with

Appendix 3 translates these areas of interest into interview questions

e.g. Managers Topic 2: what and who were the key influences on the design of
the assessment?
e.g. Supervisors Topic 4: have you ever had to fail a student? Tell me about it.

I adopted a flexible approach open to changes of wording and sequence, to
allow for the participants to speak freely about the topic in their own terms
(Robson, 2011) and to explore themes as they emerged. Rather than a process
of 'mining' for information, Brinkmann and Kvale (2015) describe the interview
as one of 'travelling' with the interviewee, being curious and sensitive and open
to the unexpected. My interviewing method moved backwards and forwards
between these two metaphors, as time restraints impinged or unexpected
responses presented. My experience was that this was difficult and I will return
to this at the end of this chapter in my reflections.

3.4.6 Conducting the Interviews

Interviews were audio recorded and took place within the institution except for
the two on Skype. I kept research notes to describe the settings of the research,
keep track of the interview schedule, record comments made outside of the
recorded interviews and record my observations and feelings as I conducted
the research. Greene (2014a) recommends the use of a research diary in order
to maintain analytic distance. My experience suggests that diaries also have a
role in reflecting on ones' feelings about the interviews as it is through this that
shifts in perception can be noted.

3.4.7 The 'partial' insider researcher

The position and perspective of the researcher in relation to the practice I was researching was important to deconstruct as it is acknowledged, in a qualitative and interpretivist method, that meaning is co-constructed. My managerial perspective as course leader and personal experience as a supervisor gave me an 'insider' understanding of the issues that WBA presents to courses teaching CM. My sharing of a common identity as a practitioner within the CM community along with my knowledge of the teaching of CM had two advantages. On one side it allowed me to build rapport with the participants quickly through sharing a common language (Drake and Heath, 2008; Givati, 2012) and also attuned me to the nuances that the participants were revealing and thus provided new insights (Sikes and Potts, 2008). My similar world-view also had its drawbacks, which will be discussed below. I was also an 'outsider', not as a passive onlooker, but due to my different cultural, historical and political experience and my more distant relationship to the institutions I was researching, especially in the USA and Australia. I was also aware that I was immersed in a research practice world (Kemmis, 2009). This outside and researcher position created the 'strangeness' needed to help me acknowledge my taken-for-granted assumptions about WBA (Kuper, Whitehead and Hodges, 2013). Greene (2014b) describes how one can be an insider as being in the same cultural community and at the same time an outsider from being a social stranger, what Chavez (2008) would call a 'partial insider'. Greene (2014b) and Kemmis (2009) recommend that it is the closeness of the insider along with the distance of the outsider that makes visible the various contexts of the research and the range of discourses at work, what is excluded as well as what is included.

3.5 Data Analysis

While there was a linear process to the collection of the data, the analysis and interpretation was iterative as the conceptual focus of the research evolved with each aspect triangulating to the final purpose to reach my aims. After Flick (2004) and Robson (2011) I am using triangulation here to mean combining data drawn from different sources and people but also looking at the data from

the different theoretical perspectives of CBME and practice theory.

My approach to data analysis comes broadly under the category of thematic content analysis as described by Miles, Huberman and Saldana (2014), a method that can be used for documents and also interviews. According to Braun and Clarke (2006, p. 79) it is 'a method for identifying, analysing and reporting patterns (themes) within data'. Sandelowski and Leeman (2012) define a theme as a coherent integration of the disparate pieces of data into a meaningful and detailed account in relation to the research questions and is a way to maintain internal validity (Yin, 2014). Values, attitudes, beliefs and meanings were illuminated through my developing interpretation of the texts (documents and transcribed interviews) and through an understanding of the context in which they had emerged (Brinkmann and Kvale, 2015). While each data source was treated independently in the descriptive phase the overall aim of the research required an amalgamation of the data and findings (Baxter and Jack, 2008) leading to an interpretation, my interpretation of the phenomenon of WBA.

3.5.1 Analytical procedure

The analytical procedure I have adopted following Boyatzis (1998) was an integration of a theory driven (a-priori or deductive) approach and a data-driven (post-priori or inductive) approach. The initial review of the literature on WBA, gave me the theoretical concepts and also sensitivity to guide the question formation, collection and also analysis of the data (Wengraf, 2001). Coming to analytic generalisations was not just about supporting or rejecting my theoretical concepts from the literature but how the theories manifested in order to come to a generalization at a higher level than the specific case. Creswell (2007) proposes that an inductive approach is more conducive to understanding the participants' 'life worlds', opening up unanticipated patterns that can lead to new hypotheses, digging beneath the surface of the participants' experiences. Braun and Clarke (2006) remind us however that themes do not emerge as if they reside in the data. This was my interpretation that was emerging, a co-construction between my developing understanding of the research topic and how I interpreted what I was hearing/reading. As Boyatzis (1998, p. 8)

describes, the conducting of qualitative research involves 'emotional, value-laden, and theoretical preconceptions, preferences and world views' on the researchers' behalf that impact throughout the research process.

3.5.2 Document analysis

A descriptive analysis of the accreditation documents was undertaken by first producing a cross-case table that captured the content of the documents that might impinge on assessment as presented in Table 8 First Analysis Codes below. The codes were condensed into themes that form subheadings in chapter 4.

For the institutional documents (course and clinic handbooks and assessment instruments) my initial analysis was also descriptive and comparative. I also included a description of the clinics and their operation at this stage. Summaries and examples of the codes are presented in Appendix 5. The themes, which formed the subheadings in Chapter 5, were merged with the analysis of the interviews with managers.

Table 8 Accreditation documents codes (examples in Appendix 4)

First Analysis Codes	Second Analysis Codes
Name of accreditation board	Philosophy
Authority/ Statutory regulation	Community
Membership of accrediting body	Communication
Documents	Public
Length of programme	Accountability
Standards	Management
Core curriculum / Content	Resources
Clinical Hours	Evaluation
Nature of clinic/resources	Mapping
Competencies/learning outcomes	Hours – inputs & outputs
Theory and practice	Supervisors training
Knowledge/skills/aptitudes	Teacher qualifications
Assessment and Fairness	Variety of supervisors
Measurement	
Themes	
alignment with regulators	communication
inputs and outputs	teachers and teaching
atomisation and measurement	resources
assessors and fairness	the (lost) discourse of CM

3.5.3 Comment on first phase of analysis

Some researchers regard documents as 'fixed' and 'stable' artefacts adding credibility to qualitative research (Denscombe, 2010). However, from a socio-interpretivist perspective, they too can be seen as reflecting the authors' biases (Yin, 2014) and politics (Trowler, 2001). As proposed by Gibson and Brown (2009) documents can offer analytic possibilities that can create insights into the practices of an organisation. After I had conducted the interviews and began to understand the implications of the discourses within, I came back to the documents and looked at them differently in the light of the interview analysis. In my second qualitative content analysis I was looking for words and phrases that expressed the socio-cultural-material discourse of the documents' authors, the training of teachers and the discourse of Chinese Medicine as in Table 8 above. The themes produced added to the sub-headings in chapters 4 and 5.

3.5.4 Interview analysis

My analysis of the interviews broadly followed the 6 phase structure described by Braun and Clarke (2006): reading and re-reading, generating initial codes, searching, reviewing and defining themes and producing the report. The need to reduce the data but keep it in context was critical to develop closeness while also maintaining a distance from the data. I found that using qualitative analysis software NVivo to code the interviews led to my focusing on the technology and forming premature coding that lost the sense of the whole what wGilbert (2002) calls the 'coding trap'. Using a simple system of coloured highlighters in the text, comments and forming tables for cross-case comparison, allowed a better movement both within and between transcribed interviews. Comparing across interviews, especially as I was often cross comparing multiple documents was unwieldy on the computer.

3.5.5 Interviews with Managers

In the first phase the reading and re-reading of the transcripts from the managers of each institution produced a set of cross case descriptions against

key issues (Appendix 6a). This alerted me to the similarities and differences across institutions. At the same time I was developing analytic codes. The codes developed from the interviews and my observations in the clinics along with an example of the explanation of some of the codes and an extract from an interview showing the coding trail appear in Appendix 6b.

3.5.6 Interviews with Supervisors

In order to compare and reduce the data I combined the supervisors' responses to each interview question before coding. I then wrote a case description for each supervisor (Appendix 7a), highlighting the ideas and quotes that were suggesting a-priori codes and also starting to illuminate post-priori. While Yin (2014) suggests that case description is used when either deductive or inductive analysis has not worked, Braun and Clarke (2006) recommend writing rich individual case reports as accurately as possible as a useful method to militate against premature interpretation.

The following Table 9 are the codes developed in relation to the supervisors. These codes, including quotes from the interviews, form Appendix 7b

Table 9 Supervisor interviews Codes

Training
Learning to supervise
Teaching methods
Supervisors view of self
Relationship with students
Issues of supervision
Protecting patients
View of students
Unique views
Community
Time
Assessment – failure to fail and grade inflation
Assessment – marking and grading
Assessment instrument
Didactic vs clinic learning
Formative assessment
Expectations of students

Similarly for the interviews with the managers, the inductive stage of data analysis started as I was conducting the interviews as the clash between what my expectations from the literature and my own personal experience had been and what I was hearing alerted my attention. What was disturbing was as much about what was being revealed as what was not, what Goffman in Friedman (2010, p. 6) calls 'attention and disattention' due to the selective focus on my part and also in an attempt to understand the information coming from the participants. By the time I came to code the interviews I already had inductive themes developing. This is an iterative process and not sequential involving a merger of discovery and interpretation (Stake, 2005).

3.6 Ethical Considerations

I followed the British Educational Research Association (BERA, 2018) guidelines for ethical research including 'ongoing' informed consent, anonymity and respect for participants including guidelines with storage and transcription of interviews and minimization of anxiety through being sensitive to interviewees.

It was important that the purpose and method of the study was made clear so as not to mislead participants and this did lead to one institution, and perhaps others in the USA, declining to participate. The purpose, interview procedure including audio-taping and transcribing of interview data, confidentiality and informed consent was included in the participant information sheets given prior to collection of data held within the institution and prior to interviews. Signed informed consent was gained from the course managers and the individual supervisors. The consent form gave participants the right to withdraw from the study up to 3 months after the date of the interview and no one asked for their information to be withdrawn.

Accreditation documents were freely available in the public domain and so anonymity was not required. In relation to the institutions, because of the smallness of the CM community within each country, anonymity has been difficult to maintain and so I have been careful, as much as possible, to not

include any identifying features. As institutional documents are in the public domain I have not used direct quotes that could be tracked to source but paraphrased while maintaining meaning for the study. Anonymity of individual managers and supervisors has been maintained by using pseudonyms. However, the nature of a cross case analysis has needed some level of identification of country and thus potentially the institution, in order to show the development of patterns and possible relationships.

The key to interviewing practitioners in their professional role is that it relies on the researchers' integrity (Brinkmann and Kvale, 2015) including being sensitive to any ethical issues that might arise. Integrity was apparent early in the research as I made choices of the institutions in the UK to investigate. Integrity has also been maintained through transparency of procedures, clear descriptions of the evidence chain, triangulation, alternative interpretations and care with causal explanations.

Brinkmann and Kvale (2015) present a series of issues in relation to ethics. Within most interviews there is a power asymmetry as the interviewer has control of the questions and ultimately the interpretation and needs to be sensitive to the interviewee. They describe the need for the interviewer to create the conditions where the interviewees will feel both free and safe enough to reveal private understandings that are being recorded for later use in the public domain. Mostly participants were open and generous in their answers. The interview of 9 minutes was an example of a decision I made on the basis of how informed consent might have been misunderstood. This one supervisor, while signing the consent form, seemed extremely reluctant to be interviewed and, I felt, wary of my intentions. It was also taking place in their lunch break between a busy morning and afternoon shift and they did not want to schedule at another time. I called a halt to the interview as I felt they may have misunderstood the aim of the interview and that I did not have 'ongoing' informed consent.

All digital recordings were transferred onto my personal computer that was password protected and the original recording devices wiped. A verbatim

transcription was made of the interviews. Following the Data Protection Act 1998, all documents, recordings and transcribed data have been stored in a secure and lockable filing cabinet in my home. Once the interviews were transcribed the recordings were wiped. All reference to places and names that might identify participants have been blanked on the transcripts.

3.7 Reflections on Insider-Outsider Research

For Holstein and Gubrium (2004) the interview should be seen as 'active' and recommend paying as much attention to the process of the interview as to the product as it is in the interview that meaning is actively assembled. As background to my interviewing method I adopted what Dahlberg and Dahlberg (2004, p. 272) call ' bridling' as opposed to 'bracketing', that is kept in check my perspectives and interpretations as opposed to keeping them out (Lillrank, 2012). My experience was that this was difficult and in the initial interviews I note my lack of 'bridling'. As an 'insider' researcher, a sharing of information on my own institution or my personal experience did allow for a development of rapport but it is the revealing of opinions that can lead interviewees. Just as practice theory challenged the constructs of CBME, so too is the researcher not an isolated, objective and neutral person but comes to the interview with a personal context and needs to be able to reflect-in-action.

My sense of being an 'outsider' was also marked as I had anticipated a familiarity due to my extensive experience in education of Chinese medicine. My choice to not study my own institution but to seek out other institutions and those in other countries was an important decision. If the culture and context is different then it is likely that the enactment of assessment will be different and the information I was getting along with my reaction to the information was helping me to see differently. My research note reads:

> *I do not have the questions clear in my head and so needed to keep looking at my notes; but that is because what they* [the supervisors] *do does not match what I expected and so my questions are irrelevant; this is proving difficult. 27/05/'16.*

The initial interviews I found the most difficult as some of the questions I had decided to ask did not seem relevant to the experiences the supervisors were presenting. For example, my aim in asking about failing students was to open a dialogue on the difficulties of moving from teacher to assessor and on the role of personal judgement in assessment; but failing students was not an issue for most of the supervisors. However, at the same time, it was the gulf between the interviewees' experience and mine that was revealing new knowledge. This recognition was illuminating and I noticed my reactions going from confusion as to what I was hearing to a beginning understanding of the very information I was trying to find out about. Goffman's (1986) view is that our tendency is to project our interpretive frame into the world and only shift when some incongruity forces a change. That frame-shift happened at numerous points throughout the research, both during the data collection and throughout the data analysis, as the hermeneutics of the research imposed on me. It was my own personal and institutional mind set that was influencing my interpretation (Orr, Cunliffe and Bennett, 2009). Brinkmann and Kvale (2015) suggest that in the co-construction of knowledge from interviews, ambiguities and contradictions can arise due to failure of communication. This was not failure of communication but of interpretation on my part. One of the further challenges Kvale (1994) poses is that interview research tends to be idealistic and ignores that human experience and behaviour is situated in a social and a material world. As Orr, Cunliffe and Bennett (2009) recommend, as researchers we need to consider the constraints and opportunities that 'insiderness' and 'outsiderness' present, the dynamics of the relationship with the interviewees and I would recommend turning that attention on documents as well.

A common critique of case study research is that there is a bias towards verification of the researchers' position especially as any theories have been developed out of the context of the specific research. Flyvbjerg (2011) counter argues that it is more likely that the bias will be towards falsification of any preconceived ideas, which is what I became aware of. From a socio-interpretivist perspective, however, this restraint from external imposition is debatable as the researcher herself comes to the case with assumptions and preconceptions and is constructing meaning through the very questions asked

and choice of focus as well as interpretations made. Baxter and Jack (2008) consider that a drawback of entering case based research with a conceptual framework is that it may limit inductive insights. While not putting forward predictive theories (Flyvbjerg, 2011) some theoretical concepts might be needed to form the criteria on which outcomes are to be judged and Yin (2014) proposes that even exploratory studies need statements about what is being explored and the purpose of the exploration. My experience was that my conceptual frameworks acted in tension with each other allowing for a dialectical movement within my analysis and interpretation. I was left, however, with a concern that the lens by which I viewed the practice of WBA still filtered my perception especially in whether I was able to view the institution in the UK with the same 'strangeness' that I viewed the other two, as will become apparent in the analysis and discussion.

Summary

In this chapter I have outlined my strategy for researching WBA in three CM institutions in the West. I have described how the study was conducted and highlighted some of the problems that can arise in qualitative research of this kind. My methodological aim has been to see patterns that others may not have seen, or have not attempted to, to make them as transparent as possible, not in an objective way but as we do in making a medical diagnosis, to justify what we have concluded not deductively or inductively but abductively (Moir, 2014), that is aiming towards the best logical inference and explanation for the phenomenon concerned. By moving outside my own life-world, to butt up against one in which I was an outsider, allowed for the 'strangeness' needed to gain insight into how WBA was being enacted.

Presentation of Analysis in Discussion Chapters

In the next three chapters 4, 5 and 6, using the analytical approach outlined above, I tell the stories of assessment from the point of view of the different stakeholders. I present the final themes as the subheadings in Chapters 4, 5 and 6 combining an analysis of my empirical data and my interpretation

(Brinkmann and Kvale, 2015; Robson, 2011). Chapter 4 will consider the accreditation documents of the three countries in relation to their requirements for institutions teaching CM. Chapter 5 will analyse and discuss the institutional documents combined with the interviews of the course managers. Chapter 6 will present the perspective of the clinical supervisors. My field notes also provide a background to the later two chapters. Through the similarities and differences between the different stakeholders and between institutions, I will explore how each was enacting CBME and how WBA was being enabled or constrained. Chapter 7 brings together the entwinement of the three data sets through a practice theory perspective.

Chapter Four: Accreditors – standardization and accountability

Accreditation processes have a primary aim in maintaining standards in professional education for the sake of students, members of the profession, and patients. Accreditors also have a role as guardians of the profession and of providing guidance and support for institutions. In this chapter I will show the forces that shape their policies and procedures as their quality assurance role is integrated with their guidance and as their role in maintaining status for the profession is integrated with their duty to uphold the professional culture.

4.1 Alignment with Regulators

Each country had different regulatory processes that acted explicitly or implicitly on the decisions accreditors were taking in relation to their standards and requirements for institutions (see previously Section 3.4.3).

In the USA, because of the licensing requirements for practitioners, all institutions in the USA had to meet accreditation standards. In Australia accreditation was carried out by the Chinese Medicine Board of Australia (CMBA). Institutions teaching Chinese Medicine were also governed by requirements of the Tertiary Education Quality and Standards Agency (TEQSA). In contrast in the UK, there was no statutory regulation for CM and so accreditation was a choice for institutions.

In spite of these differences in statutory regulations, the accreditation handbooks were very similar in the standards they set for institutions and similar to other health care professions including those related to WBA. Either embedded in the documents or separately were the professional competencies/capabilities/ standards[8] that institutions were expected to adhere

---

[8] The term standards is used in relation to what accreditation boards require of institutions but is also used in some documents and in institutions synonymous with competencies. As much as possible I will use standards in relation to accreditation requirements and competencies or (learning) outcomes when referring to discipline specific knowledge, skills and aptitudes.

to. CBME was the curricular model chosen. While not legally required, the BAAB, an independent organisation, had brought its Standards of Education and Training for Acupuncture Programmes (SETAPs) (BAAB, 2016) in line with requirements of the Health and Care Professions Council (HCPC). In the USA the alignment with other health professions was shown by the ACAOM (2019) publication of their professional competencies that had moved away from the previous structure that was constructed around a CM consultation and treatment (ACAOM, 2016), to a structure adopting much of the language used by the Accreditation Council for Graduate Medical Education (ACGME). The struggle for parity with the dominant health professions as a means to enhance professional status aligns with the history of CM in the PRC as discussed in the introduction. Striving for status and statute through a process of professionalization (Saks, 2015), has also been the driving force behind the CM profession in all three Western countries for the last 30 years. While the USA and Australia have managed to gain control of title and thus a level of government support, in the UK the current statutory health professions continue to block access to resources for CM practitioners through professional closure (Saks, 2001).

The developmental and supportive role of accreditation was also presented in the opening pages of ACAOM and BAAB. Their purpose was stated as: 'to encourage diversity, experimentation and innovation within the boundaries of generally accepted standards and guidelines of academic quality' (ACAOM, 2016, p. 1; BAAB, 2016, p. 5). That the CMBA did not define their purpose in this way. I propose, was the result of the further level of statutory regulation from the AHPRA, their role more focused on accountability.

4.2 Inputs and Outputs

The absorption of CBME as the primary curricular framework was displayed in all three countries as shown by the detailed and atomized descriptions of student 'outputs' in the form of competencies or learning outcomes which I will discuss further. At the same time, however, the USA and the UK also stipulated 'inputs' in the form of hours of teaching and learning.

(ACAOM, 2016, p. 34)
The minimum length of the professional acupuncture curriculum must be at least three academic years (a minimum of 105 semester credits or 1905 hours)

- 47 semester credits (705 hours) in Oriental medical theory, diagnosis and treatment techniques in acupuncture and related studies
- 22 semester credits (660 hours) in clinical training.

(BAAB, 2016, p. 29)
… the expectation is that programmes will be no less than three years full time or part-time equivalent with a notional total of 3,600 hours. Two-thirds of this is likely to be independent study but a minimum of 1,200 hours should be structured time when students will either be in direct contact with a member of staff of the institution or working on aspects of the curriculum in a structured and interactive way. Most importantly, no fewer than 400 of those hours should be related to clinical practice in the direct care of patients.

In these extracts we can see these two accreditors clearly specifying the hours teaching institutions were to provide including clinical hours. The breakdown into hours comes from the pre-CBME concept of education discussed in Chapter 2.1 in which 'inputs' (content and hours) determined students learning. Lucey, Thibault and ten Cate (2018) report that many medical schools are still applying a 'time-based, competency-variable paradigm' (p.S1) concluding that CBME has not been taken up in its entirety by medical schools even as it is being promoted by accreditors. In these examples it was the accreditors who were still holding onto a time-based curricular model along with CBME. In contrast, in Australia, the CMBA did not stipulate hours. What they required, which is more in line with the ethos of CBME, was that the institutions justify their contact time. The accreditor seemed to be imposing a new model of accreditation that derived from TEQSA requirements. For the USA and UK, the full manifestation of CBME was less in evidence, showing external regulators, for now, having less impact.

In the USA, the 'input' view of achievement was even further defined illustrating the differentiation between didactic and workplace learning as presented earlier.

(ACAOM, 2016, p. 35)
One semester credit is granted:

- For each 15 hours of classroom contact plus appropriate outside preparation or the equivalent;

- For each 30 hours of supervised laboratory or clinical instruction plus appropriate outside preparation.

This extract reinforced the time-based approach of this accreditor to calculating if a programme was meeting required standards. However, there was a further implication in this extract. An hour spent in a classroom context was seen to have a higher credit worthiness (15 hours per credit) than an hour in a clinical context (30 hours per credit). This differentiation manifested, I argue, the acquisition concept of learning as more effective than participation (Sfard, 1998) and implied that authentic settings were being seen as less valuable in relation to learning. An update of the ACAOM document for 2019 no longer has this definition and one has to consider that the ethos of CBME is being followed more closely with outputs being the quality assurance measure. The concept of hours as indicators of learning is a theme that will recur and was manifest in all three institutions.

4.3 Atomization and Measurement

Outputs in the form of competencies broken down into sub-competencies were listed in all accreditation documents with requirements for how institutions were to measure them. The formulations in these documents presented three key steps in the characterisation of assessment of learning in CM namely that what institutions were required to assess was the 'performance' of the students, that students must have 'demonstrated' through this performance that they met each of the learning outcomes and that these would be 'measured'. Further, the accreditors required that institutions' learning outcomes be mapped onto the prescribed professional competencies to ensure all were covered. The assessment of competence in these instructions meets the instrumental-technical expression of CBME that is the measurement of a student's performance to be displayed in a standardized psychometric format. 'Diversity, experimentation and innovation' was not evident in their approach to assessment.

The deconstruction I give below demonstrates the difficulty that can be incurred when each sub-competency needs to be measured independently rather than

as integrated competencies. This example is similar to those found in the sub-units of the other accreditor's documents. The (CMBA, 2016, p. 49) integrated competency,

> 6.6 to safely, skilfully and effectively practise acupuncture according to the relevant Chinese medicine theories and principles of treatment.

was broken down into 14 subunits of which 2 are illustrated below.

> 6.6.3 perform a complete and <u>accurate</u> problem-focused physical examination, including tongue and pulse examinations as used in Chinese medicine

> 6.6.4 <u>interpret</u> and <u>analyze</u> the information gathered during interview and physical examination, using their knowledge, including the application of differentiation of diseases (*bianbing*) and patterns (*bianzheng*) in Chinese medicine, to diagnose the presenting health issue

What these extracts illustrated was a separation of the performance of the physical examination as distinct from the interpretation and analysis. Not a 'dictionary' to narrate competence (Tekian *et al.*, 2015) these were to be measured independently. If competence must be inferred from performance, an 'accurate' examination, I argue, relies on a concurrent integration of interpretation. The original integrated competency 6.6, allowed for the more complex level of interpretation that meets Hager's (2017) critique of the problems incurred with atomization. While in other sections of the documentation, the accreditors called for an integration of theory and practice within an authentic setting, it was CBME's performance, psychometric and product discourse of competence (Hodges, 2012) that was the direction being taken by accreditors. Further, as argued already, the requirements of atomization sets limits on the curriculum to a defined content, disregarding of the holism required in the workplace. Givati (2012) has questioned how one can both standardize and, I add, atomize competencies and maintain a unique ethos true to the philosophy of Chinese medicine.

4.4 Assessment and Fairness

The quality assurance requirement through measurement was displayed in the call for 'fairness', 'objectivity', 'validity and reliability' in assessment decisions.

(CMBA, 2016b, p. 32)
describe the quality assurance processes you use to ensure the continued <u>reliability</u> and <u>validity</u> of your assessment methods.

(BAAB, 2016, p.45)
The measurement of student performance must be <u>objective</u> and ensure fitness to practise. The <u>criteria</u> used for assessment at all stages of learning contribute to the necessary <u>objectivity</u> for any <u>fair</u> marking system.

Maintenance of standards through the validity and reliability of assessment would be gained by the use of agreed standardized competencies, the main purpose of CBME. However these extracts were also accompanied by a quite different way this was to be achieved.

(BAAB, 2016, p. 45)
Another useful strategy is double marking or <u>moderation</u>. Assessment practices should be clearly set down in a marking policy.

CMBA (2016b) asked institutions to:

describe your assessment processes and procedures, including <u>moderation</u> procedures. (p.32)

which was there to provide:

comparability of standards of student performance across, for example, different markers, locations, subjects, providers and/or courses of study. (p.61)

Moderation, I have argued, rather than promoting a uniform interpretation in order to balance assessor bias, allows for differences in interpretation to emerge (Govaerts, van der Vleuten and Holmboe, 2019; Wilkerson and Doyle, 2011). It is the standardization of the mapping process required to provide objectivity, that assumes a clarity of meaning that can be manifest across an institution's unique objectives and goals, different conceptualisations and different communities and context. The ethos behind why accreditors thought moderation might be necessary came from their desire for objectivity.

(BAAB, 2016, p. 45)
Achieving a measure of objectivity in assessing clinical work is particularly challenging, and includes recognition of the possible influence of <u>personal values and beliefs</u>.

Similarly to the other accreditors, this statement demonstrated the concern for the accreditors that the assessors' 'personal values and beliefs' that they

brought to their role, their bias, might challenge the objectivity required for establishing fairness. The BAAB seemed to be calling for a check on any local translation of competencies to reduce assessor error as opposed to celebrating the valid idiosyncrasy of each assessors' decisions as discussed in the literature review. Objectivity would be realised through calibrating the assessors to the assessment instrument.

4.5 Communication and the Input of Supervisors

For moderation to be achieved would require communication amongst supervisors and between managers and supervisors. The requirement for communication, however, was limited in the documents. ACAOM (2017, p. 26) required that staff should meet 'several times a year to consider educational policies and issues' and CMBA (2016, p. 12) to demonstrate how 'clinical supervisors were included in the quality assurance systems'. In the BAAB there was more detail given around the role of clinical educators.

> (BAAB, 2016, p. 38)
> Communication between clinical supervisors and their peers and managers responsible for curriculum development is essential and formal structures for this need to be in place. Clinical supervisors also need to be involved in curriculum design, curriculum changes and evaluation of teaching and learning as well as in the setting up and the implementation of assessment strategies for practice experience.

This example illustrated a recognition of the need for more than participation but that communication between supervisors themselves and also with other staff members should be fostered. Clinical supervisors were not identified in this way in the other accreditation documents. If sharing assessment experiences and moderation processes were critical for developing shared mental models around learning outcomes, two of the accreditors had little to say as to how this was to happen.

Drawing on Kennedy *et al's* (2015, p. 5) proposal that it is the 'talents, interests and capacities of educators' that will be 'pivotal' in determining outcomes for students, my focus turned to what accreditation standards or guidance there was for the training of supervisors. All accreditors required that teachers were appropriately qualified in their discipline but there were much lighter

recommendations on the teachers as educators.

> (ACAOM, 2016, p. 25)
> The general education, the professional education, the teaching experience and the practical professional experience must be appropriate for the subject area taught.

> (CMBA, 2016b, p. 30)
> have an understanding of pedagogical and/or adult learning principles relevant to the student cohort being taught.

> (BAAB, 2016, p. 19)
> All teachers of acupuncture need to see themselves both as professional acupuncturists and as professional teachers in higher education.

While these extracts illustrated that maintaining pedagogical and discipline skills were expected, the emphasis of the documents erred more towards the discipline. ACAOM was the least robust in its requirement for teacher training referring instead to 'teaching experience' and the CMBA called for 'understanding' only. In the UK accreditation document there was a more equal emphasis on training in CM and in education. What I propose, is that in these examples it is the accreditors' role in upholding the standards of the profession of CM that are more the focus, not the discipline of teaching. The tension created, however, is that a lack of pedagogical training can lead to supervisors having a simplistic understanding of CBME. This lack of teaching qualifications in professional schools is reported elsewhere: Tekian and Norcini (2016) in the USA and Bussey (2019) in relation to teaching qualification for undergraduate medical supervisors in the UK.

In the USA and UK the accreditors called for a 'variety' of supervisors in terms of numbers and perspective.

> (ACAOM, 2016, p. 37)
> The number of clinical supervisors must be sufficient to ensure effective instruction of and safe practice by interns. Student interns must receive training from a variety of clinical faculty members.

> (BAAB, 2016, p. 37)
> Students need to learn from more than one clinical supervisor in any given year of the programme. Offering more than one perspective on each student's clinical progression will enhance their range of experience and add to the validity of assessment of practice.

This call for variety is implicitly in the USA and explicitly in the UK, a valuing of differences in practice that can come from supervisors and the benefits this may confer on students' learning. In these examples one can see, as in the need for moderation, a move away from the measurement of students' performance in relation to the assessment outcomes as the determiners of competence to the role of supervisors. In the BAAB these different perspectives were also recognised as supporting the validity of assessment. Combined with the statements above, the complexities of WBA were manifest, with an accountability purpose of objectivity and inter-rater reliability in tension with the authenticity that comes from the supervisors steeped in their professional culture.

4.6 Resources

In relation to the authentic environment, a key construct of CBME, all three accreditors had standards around resources for clinical training calling for 'clinic space with sufficient equipment and facilities' (ACAOM p.29) and evidence of the 'development of a viable teaching clinic' (BAAB p.57). Standard 2.7 of the CMBA (2016b, p. 18),

> providing the teaching and learning equipment and devices, as well as the equipment relevant to clinical practice, necessary for Chinese medicine students to achieve the program learning outcomes

while describing the clinical artefacts required, added in a linking of those artefacts to learning outcomes. While this could be taken at face value for a practice-based programme, it can be read as introducing a performativity agenda. This link of resources to learning outcomes can also be seen in the later iteration of the ACAOM (2019) standards.

> (ACAOM, 2019, p. 5)
> Program facilities must be sufficient to meet the program's statement of purpose, and to support its faculty, students, goals and learning outcomes.

The earlier iteration of ACAOM 2016 (p46) linked the requirement of learning resources and equipment to the educational programme in general rather than tying it to learning outcomes. The change implied that the accreditation process

would now measure resources against student outcomes as opposed to the process of education. This example hints towards concerns that education has shifted its focus from pedagogy to accountability (Ball, 2003; McEwen and Trede, 2014; Orrell and Higgs, 2012).

4.7 The Lost Discourse of Chinese Medicine

Accreditation boards have to enact their standards over multiple and diverse institutions with quite different communities and locations and the accreditation documents called for a consideration of the wider context in which institutions were placed.

> (ACAOM, 2016, p. 5)
> The statement of purpose and goals should reflect the <u>purpose</u> for which the program was founded, the <u>philosophies</u> it represents, the <u>community</u> in which it is located, the constituencies it serves, the needs -- <u>social, cultural and material</u> -- of its community and clientele, and the institution's resources -- <u>human, physical, and financial</u>.

> (CMBA, 2016b, p. 9)
> Do you think the program is well-regarded by the broader Chinese medicine professional <u>community</u> and health sector?

> (BAAB, 2016, p. 23) required documents to show the
> … educational <u>philosophy</u> and the <u>relationship</u> of this to healthcare in the country and the local <u>community</u>

These statements show in different ways a recognition of the contextual influences on programmes and institutions, that they might have a unique philosophical ethos and be part of a wider community. However, the documents lacked any philosophical or cultural discourse of CM. This mirrors a concern noted by Collini (2017) and Schwandt (2005) in relation to HE in general. They have seen an increased precedence being given to the accountability language of NPM. Bleakley, Brice and Bligh (2008) confirmed by the research of Ho *et al.* (2011) and Ho *et al.* (2014), as discussed in the literature review, warn that the very use of the language of core competencies and standardization, promotes a particular set of values, particularly Western values. They propose that it is the way knowledge is organized and operationalized that stifles both cultural and

political differences - an homogenization of practice reminiscent of the situation in the PRC as discussed in the introduction.

The competencies in the USA and Australia also lacked some of the language specific to CM. For example, present in the ACAOM 2016 document, the word '*qi*' was not present in the 2019, nor in the competencies of the CMBA. 'Body-mind-spirit' and 'holism[9]' were also terms not given expression. This lack of a CM discourse aligns with Givati's (2015) critique showing accreditors to be leaning more towards the expectations of outside regulators and an accountability discourse rather than the profession of CM.

The exception to this was the professional competencies of the BAcC that did include these expressions. The BAAB was also most constrained in its use of an instrumental-technical language and was probably the closest to displaying some concession to a discourse more descriptive than accountable. They listed a set of values they wished to see in institutions that included:

> (BAAB, 2016, p. 7)
> practice-led: rooted in the <u>artistry</u> and skills of acupuncture as an empirical and practical professional activity informed by theory and <u>creative of theory</u>, and recognizing that, as acupuncture is a practice-based profession, acupuncture teachers will normally remain engaged in practice

Expressions such as 'artistry' and 'creative of theory' are less easily contained by metrics. Skills that are primarily intuitive and artistic and the affective elements of professionalism is what challenges all professions when an outcomes framework dominates. That this language was being employed, I propose, was probably due to the profession being most removed from any

---

[9] There is not space here to present in detail the historical arguments and problematic of the expression 'holism' and its entanglement with CM and I refer the reader to Scheid (2016). Holism is a term that explains but does not define CM. Not an expression found within the classical canons of CM, the concept of holism in CM is only 100 years old (Scheid 2016). It emerged as a dialectical tool that helped to differentiate CM from the reductionist scientific orthodoxy of Western medicine (Scheid and Karchmer, 2016). The expression 'holism' explains a fundamental construct of CM that maintains a unifying perspective on body-mind-spirit, and a personalized approach to medicine that sees each individual as unique in their presentation of signs and symptoms.

regulatory requirements due to the legal situation of Chinese medicine in the UK as described in section 3.4.3.

Summary

This analysis has shown that it is the instrumental-technical discourse of CBME as shown in the drive for accountability as outputs and measurement that predominated in the accreditation documents of the three countries. This follows a global trend as discussed by Fishbain, Danon and Nissanholz-Gannot (2019) in relation to medical education accreditation systems.

Hordern (2014, p. 173) considers that accreditors can either lead or be led by change, facilitating a new conceptualization of the profession or complying with pressures from other dominant professions. My conclusion was that the striving for external legitimacy and social status was having a major influence on the accreditation documents. While meeting the accountability purpose of accreditation, standardization and codification was setting limits on the heterogeneity and multiplicity of the practice of CM. That the BAAB in the UK was employing more of a language allowing for 'indeterminate, experience-based' and 'artistic skills' (Givati, 2015, p. 47) was due, I argue, to the lessening of ties to external regulators. Boyd *et al.* (2018, p. 46) suggest that CBME is a 'social construction embedded in its political and historical context'.

The socio-constructivist approach of CBME was shown in the instruction for WBA to take place in authentic settings and all three accreditors had detailed standards and descriptions on the resourcing of clinical teaching along with the requirement for experienced practitioners of CM as supervisors. However, the construct of measurement also dominated the role of the supervisors as assessors. While the personal beliefs that can ensue within authentic settings can be seen as valid interpretations, the accreditors were more concerned that this would challenge the objectivity of assessment. And yet they had little to say on the training and communication that is required in order to understand

CBME and how shared mental models were to be developed to aid a fair assessment process.

Hager and Gonczi (1996) suggest that how professional accrediting bodies conceptualize competence is going to have an impact on how competence is assessed at the institutional level. However, Edwards (2010) considers that how institutions react to accreditation standards is unlikely to be a mechanical reaction to pre-established conditions. In the next chapter I will consider how institutions were managing the different constructs of CBME.

Chapter Five: Institutions – leaning out and leaning in

This chapter will consider data from the handbooks in the institutions and the interviews with the managers. As presented in my introduction, teaching and assessment methods are not spontaneous but are adopted and adapted by institutions according to particular and dominant discourses of pedagogy, curricular policies and external stakeholders (Hodges and Lingard, 2012b). In exploring how standardization was influencing the enactment of WBA, I also investigated whether a place had been found for the subjective, interpretive, inconsistent and intuitive nature of professional work (Schwandt, 2005) that is manifest in an authentic environment. How were the institutions managing the need to measure a student's achievement of multiple learning outcomes and the place of supervisors' judgements and still maintain inter-rater reliability? Did tensions arise from the absence of a CM discourse in the accreditation competencies? The following analysis will link back to the previous chapter as I consider the alignment of the institutions with the accreditation documents and at times forward to the story yet to be told of the assessors.

5.1 Assessment Instruments and Measurement

In all three institutions, accreditation requirements were manifest in the detailed articulations of learning outcomes embedded in assessment instruments and described by the managers as:

> *all of the parts of the curriculum are supposed to connect with the course LO's and the course LO's connect with the programme LO's. We also blueprint all of our assessment items so all faculty are required at mid and end exams to say which questions are aligned with which LO. (Mike)*

> *so the course learning outcomes address the whole programme and the subject learning outcomes the individual subjects by themselves. So they are constantly being tweaked. (Mark)*

The mapping of learning outcomes from the course to the subjects and onto the assessments was consistent with the discourse of the standardization of competencies promoted in CBME.

All WBA instruments were unique to each institution and contained multiple learning outcomes. In one format 14 competencies were broken down into 97 outcomes requiring not a mark but a rating of 'competent' or 'not competent'; written feedback on each outcome was expected. Another format listed 20 outcomes in relation to the consultation and treatment to be annotated with a numerical mark for each. In the third, 5 categories were broken down into 25 learning outcomes with a grade for each to be annotated. In two instruments the assessor would annotate each learning outcome against 5 grades from not-competent/fail through to excellent/70-100%', which, while requiring a greater level of judgement on the part of the assessor, would make it easier for the conversion to a mark. As discussed, the literature challenges the objectivity of assessing multiple learning outcomes (Klamen *et al.*, 2016; ten Cate and Scheele, 2007), however, all instruments aligned with this atomized construct.

There were further complicating issues related to the format of the instruments. In one institution the way the outcomes were written, and with no qualifying criteria, seemed to be asking for a yes/no answer and yet a differentiated mark was expected.

Did student effectively elicit a thorough medical history?   Mark ___

The format of the instrument with its requirement for a numerical grade seemed to direct any assessor to the allocation of full marks. In another a final mark was expected of the supervisor yet each outcome only called for a competent/not competent decision and there was no space for a mark. My analysis of most of the assessment formats, especially when there were no differentiating criteria, led me to question the validity and reliability of any differentiating mark given. The quantified symbol created a level of certainty that belied the lack of objectivity in how the mark was to be reached (Yorke, 2011a). The formats would tend, I argue, towards the allocation of an overall 'A' grade, an issue that I will follow up below. In support of my argument, in all the instruments all learning outcomes seemed to hold equal value and there was no information on whether there was compensation across them or whether all had to be passed.

In only one institution were there qualifying criteria for each learning outcome to

help differentiate the grades for the assessor (and student) and further this institution supported moderation.

> *So for that grading, we <u>calibrate</u> it amongst all the supervisors. So it's like calibrating and <u>double marking</u> merged together. So we have an assessment meeting amongst all the supervisors to make sure we don't have <u>bias</u> of any individual and we calibrate the marks between us, so if one deserves 75, what's the difference between another person getting 73. (May)*

This description of their 'calibration' process (something called for by the accreditor), and that the assessment instruments had detailed criteria showed a more refined system for determining marks giving more credence to both the validity and reliability of the final mark. The criteria descriptors and the calibration meeting together helped to create a 'shared mental model', lessen bias and generate inter-rater reliability. A moderation process that could tease out supervisors' judgements was not present in the other institutions.

5.2 Stripping the Lifeworld of Chinese Medicine

In the generic institutional documents such as Mission Statements and Course and Clinic Handbooks, a contrasting discourse appeared to that conveyed in the assessment briefs and learning outcomes. In the Mission Statement of one institution it said:

> To inspire, through the presentation of East Asian Medicine, a deep and lasting respect for the <u>integrity of body, mind and spirit</u>.

This statement incorporated a discourse steeped in the holistic premise of CM that has never separated body-mind-spirit. The language used in all the Clinic handbooks infused a discourse of praxis.

> *The senior clinical placement …. is a time when textbook cases become human beings with real needs, pains, and concerns. It is a time to apply both <u>compassion</u> and integrity produced by the knowledge, skills, confidence and self-esteem developed in the classroom.*

> *Integrity: The sense of <u>personal wholeness, honesty, trust</u> and consistency are important to both the patient and the student.*

> *When in the clinic, allow yourself times to be quiet and still inside. In the same way that you become quiet and attentive when listening well to*

music, or when soaking up the beauty of a landscape spend some time sensing the quality of the <u>Qi</u>, and attuning to the more subtle aspects of the treatment process…

These extracts depicted a view of medicine that focused on the relationship between the student, the patient and of praxis (right conduct) and the subjective nature of clinical work. This was in contrast to the language employed in the assessment instructions and instruments that supported an instrumental-technical discourse, devoid of praxis, of any CM philosophy or of the role of intuition. The closest descriptions of performance in relation to personal attributes were as follows:

treats patients and others with respect, altruism, ethics and morals, and sensitivity to cultural, age, gender, and disability issues.

implement ethical work practices.

deals effectively and sensitively with patients.

Values of '*personal wholeness, honesty and trust*' or '*the integrity of body, mind and spirit*', or '*sensing the quality of the Qi*' as quoted above, might be assumed under these outcomes, however, as written those values were at best, taken-for-granted. Thus, there was a contradiction appearing in the documents. In some places the richness of the medicine, the personal and professional development of the students and the embodiment of '*qi*' was illustrated, but when it came to the assessment instruments the 'lifeworld' had been stripped out as the documents yielded to the standardization and measurement discourse. The perception of Collini (2017) and Schwandt (2005) as to the increase in an accountability language, as discussed before, was in evidence.

5.3 Authentic Assessments

Because our clinic is a very good example of work-based learning, probably the best example in the university. (Mark)

This statement from one of the managers represented the clinical training provision of all three institutions and their opinion of their value in relation to meeting HE institutions' requirements for WBL. Because of the position of CM in relation to state health-care provision, all three institutions had set up their own

teaching clinics. Over their 3-4 year programmes, the students took on more and more responsibility for treating patients with the clinical environment providing the authentic experience within which students were taught and assessed on their performance in real time and place. Consultations and treatments could be anything from 30 – 60 minutes, patients paid a subsidised fee and the clinics were all busy with waiting lists. Any one supervisor worked with groups of 4-8 students supporting the treatment of from 4-12 patients at any one time. The out-patient clinics, with easy access off the street, operated for most of the year even outside the academic calendar. In the two larger institutions the clinic was quite separate from the classrooms and administrative and academic offices and for the third smaller institution it was embedded so supervisors and patients needed to pass classrooms to access the clinic. The institutions had invested considerably in their clinical facilities with Mark reporting:

> *We did an audit in 2016 and we found we had 7,760 patient contacts and it brought in 0.25 million gross.*

This statement introduced the importance of economic viability in relation to the authentic setting. With 30% of student's contact time devoted to clinic and given the staff to student ratio (stipulated by two of the accreditors), clinical teaching was one of the costliest aspects of the course. To meet accreditation requirements, all institutions had student attendance proformas to be signed off by the supervisors and students had to complete the hours (inputs) stipulated to receive their award.

The clinic handbooks had extensive descriptions of the clinical setting and expectations of students.

> They [the student] will deliver care in a supervised environment where learning can be accomplished without fear of treatment failure.

> … in a format designed to mimic the realities of a graduate practitioner's clinic

The language being employed in these extracts, was emphasizing the authentic context in which WBL, and thus assessment, would take place, a context that

leaned inwards to the community of supervisors and students working together to support patient care.

However, while 30% of student hours were being spent in clinic, the assessment of the student's clinical performance accounted for less than 30% of the final grade for all institutions. This aligned with the USA accreditation document that allocated fewer credits for a clinical hour of learning over a didactic hour. In her critique of acupuncture training, Flesch (2013) considers that canonical knowledge (including technical skills) is given more weighting than practice knowledge. A similar observation is made by Schwandt (2005) in relation to the movement of health professions into universities with canonical learning outcomes that 'intellectualizes' the work of practitioners being seen as more valid than those derived from practice. The juxtaposition of the time and expense allocated to clinic, however, led me to question if the issue was not so much the higher valuing of canonical knowledge but one of concern for the lack of objectivity inherent in WBA. I will pursue this argument below.

5.4 Failure-to-fail and Grade Inflation

All institutions required the application of psychometrics to their WBA. Rethans *et al.* (2002) remind us that the psychometric discourse is much more applicable to simulation assessments when variance is limited, as supported in the following statement.

> *I think it is one thing to be working with students and partnering them in patient care. I don't think that lends itself well to objective assessment. (Mavis)*

Mavis recognized that the inter-relationships in the clinical environment would impact on the objectivity of assessment. In contrast, in the same institution, another manager was turning to psychometricians to help improve inter-rater reliability.

> *Today I have a meeting with our clinic evaluation system, our online system, and one of the things I always wanted with our clinic evaluation system, to aggregate trends and themes, certain shifts, students, certain CLO's [course learning outcomes), supervisors. I want to look at it to see where the gaps are. (Mike)*

For Mike, it would be through an analysis of marking trends across supervisors that objectivity could be realised. It was important to note that Mike was more removed from the day-to-day running of the clinic, his role leaning more towards the outside stakeholders such as the accreditors and validators and the development of CBME within the institution. At the same time he recognised the problem of reaching inter-rater reliability.

> *And when you bring in the whole idea of reliability and validity, it is very rare we can actually intersperse inter-rater reliability. It is very challenging to do. (Mike)*

While the subjective nature of professional work was recognised in this institution, a technical-instrumental WBA was desired.

Concerns about applying a psychometric decision to WBA was also manifest in the way managers described other assessments in the programmes to those in clinic. They all described a reliance on the classroom assessments rather than the clinic assessments to fail weak students.

> *…. and they fail academic before they fail clinic. Clinic is very difficult* [to fail].
> *But they are more likely to fail on academic subjects, and practical techniques rather than actually in clinic. (Mark)*

> *What vets[10] them out is they can't pass the comprehensive exam [an MCQ]. And it is not the clinic work which is very subjective but the didactic class. (Mary)*

What Mark and Mary were voicing was that failing students would have been picked out within earlier written or skills assessments. The assumption that previous assessments would fail those students not competent to practise misses, I argue, the very essence of WBA that captures the professional, ethical and integrated work that can only apply in the workplace. However, Mary's quote highlighted that it was the subjectivity of WBA that inhibited the failing of a student in clinic. A further example of the reliance on 'evidence' was that in all three institutions, the detailed writing up of a case history of a patient who had been treated in clinic formed a more substantial part of their final 'clinical' mark.

---

[10] To vet is an American expression that means to 'veto' or rule out.

The objective evidence of a written piece of work held more value. The difficulty of applying competencies in an objective way to WBA was described further.

> *the challenge comes when a student actually challenges* [their] *grade and they want evidence for how and why they got that grade and the faculty have to recreate in their memory what they did or did not see. (Mike)*

This extract illustrated that the need for objective evidence in relation to grade decisions was impacting on the whole assessment culture. Thus the subjectivity of WBA was implicated in failure-to-fail.

That clinic was difficult to fail was given a quite pragmatic justification linked to an assessment that had a long gestation. Competence was seen as a developmental process and any 'failing' competencies would be highlighted through the day-to-day feedback given by supervisors.

> *they'll* [the supervisor] *take a student aside and they will say, you're on the way to failing this. They are quite firm at that point. And several times over the past few years I've seen that make a student turn themselves round, because they realise, I've got ten days left to nail this, otherwise I'm going to have to do the module again. That's money and time and putting off your plans. (Matt)*

This extract showed Matt linking a competent outcome for clinical work to students' motivations in terms of costs and time. Other managers linked good performance to the majority of students being mature with no managers citing any situations when a student was actually failed a WBA. The rationale for not failing students, however, was complex and will be better returned to when the supervisors' voice can be introduced.

Along with 'failing to fail' the problem of 'grade inflation' was voiced by all institutions.

> *More often than not the supervisors tend <u>to grade very liberally</u> in that regard which is OK as the supervisors are in there a lot and they know if a student is not doing, they are in there with the student watching their needling and listening to their presentations. So I mean, I would even say it should be pass-fail because essentially they are graded as a pass-fail because it is a subjective component that the <u>students get an A</u> on unless there is something really amiss. (Mike)*

This quote along with other expressions such as *"they tend to give people too good a mark"* and that WBA marks were *"skewed to the higher end"*, illustrated two features of assessment. Firstly, the managers were aware of what they called *"grade inflation"* and considered this a normal factor of working closely with students over time. Secondly, the psychometric decision expected from the supervisors and, I add, expected from the format of the assessment instruments, could not be substantiated; an *"A"* grade was indicative of competence rather than excellence.

> *so the only reason you would give an A- in the clinic is because of attendance issues which is objective or because they have not handed in the appropriate paper-work which is objective. (Mary)*

The reason for the 'A' award being the normative grade was linked to the issue of the lack of objectivity in WBA and not to any problem with the construct of CBME (Boyd *et al.*, 2018). According to Knight (2007, p. 76) measurement needs to be applied to stable and real or tangible objects and is difficult when applied to 'contested social constructs', which is what Mary was implying. Grade inflation and its other manifestation failure-to-fail, transcends national and professional boundaries and is well documented in the health professions (Cleland *et al.*, 2008; Dudek, Marks and Regehr, 2005; Hughes, Mitchell and Johnston, 2016; Robbins, Firth and Evans, 2018; Yepes-Rios *et al.*, 2016). A closer look at the possible reasons behind this will be focused on in chapter 7. What can be concluded at this stage is that it was the impact of a lack of psychometric objectivity in WBA that determined the mark being awarded.

McNamara (2013) recommends two alternative processes due to the lack of objectivity of WBA. One is that supervisors are actually applying a pass/fail and of that almost exclusively a pass and called an 'A', which was reported by managers in one institution. The other was to weight the WBA low which was similar to another institution. Any mark given, therefore, would not have a major impact on the final grade. This research showed that the process that seemed to create the most objectivity was found in the institution that had written criteria and that had a 'calibration' meeting of all supervisors to moderate final marks, that is, a qualitative approach to objectivity.

Whether WBA should be allocated a pass/fail or be graded has stakeholders conflicted (Mejicano and Bumsted, 2018; White and Fantone, 2010). Those for pass/fail suggest it supports deeper learning through intrinsic motivation while those in favour of grades want to maintain the discrimination of excellence. For the managers in this study there were varied responses.

> *Creating an illusion that it is being graded <u>encourages</u> students to do better, even if it is only an illusion it serves a purpose… and a lot of students want good grades, a lot of students want an A. It is a <u>motivator</u>. And [non] attendance at clinic can bring your grade down … so that is another motivator to keep the grade system intact. (Mavis)*

This extract illustrated that the psychometric grades given were not just there to produce an accurate measurement of a student's performance but served to encourage and motivate students. Encouragement, Yorke (2009) considers, is quite a different aim than measurement of competence and changes the way the awarding of a mark will be understood. Yorke (2011b) is also concerned that not including a WBA mark in the final degree classification will lessen the 'esteem' of WBL.

One could argue that for these institutions the low overall percentage grade allocated to WBA indicated a de-valuing of the WBA, however, this was in tension with the costs incurred by the hours students spent in clinic.

> *we have 1,000 hours of clinic, they are being observed by the faculty every day. (Mary)*

> *Oh pretty well everyone does [passes clinic] because we have such extensive training. (Mark)*

That students who had put in the hours were deemed to be competent expressed the curriculum 'input' view that valued hours spent in clinic with a supervisor as still holding power, as seen in the accreditation documents. Not withstanding the problems inherent in this approach to competence, the importance of the time spent in a learning environment, especially in an authentic setting, is still one of the main arguments against the full manifestation of CBME. Park, Hodges and Tekian (2016) suggest that within medical education the introduction of competencies has been attached to the original 'time and content-based' curricular model with no major restructuring

and the changes that CBME heralded with the implementation of outcomes, might not be so straightforward.

5.5 Fairness in Marking

As already discussed, when it is a wide range of situations (patients or conditions) that need to be measured then performance over time is recommended as the best assessment method and that is what all institutions were doing. When dealing with unwanted variations in assessments, and when judgements are highly subjective, Crossley (2016) refers to a number of possible assessment strategies to maintain fairness in marking. One strategy is to gather a wide number of different marks – in one institution a student would be summatively assessed by three different supervisors both at midpoint and at the end of each clinic block thus generating six different marks that were combined to form the final grade. Another process is for the same judge to assess all students - in another institution two supervisors worked closely together and carried out all summative assessments. A third process is moderation which was the strategy in the third - the students worked with a number of supervisors over the year with a final mark reached after moderation and discussion amongst all supervisors. What could be seen was that each institution had unique processes for maintaining fairness in their WBA, albeit perhaps unconsciously, and not considered in the accreditation documents. These strategies, along with the limited impact of the mark allocated to clinic on the student's final degree classification, the grade inflation inferred by all managers and the hours all students spent in clinic suggested that any issues students might have with the fairness of the mark were not a factor the institutions had to contend with. Although not asked directly, only one manager reported on one student challenging their WBA mark.

5.6 Political influences

As discussed, the language of the learning outcomes in all institutions could be seen to align with the standardization requirements of the accreditors. However there were other external factors that the institutions were choosing to, or

having to, implement.

One institution was owned by a corporate organisation with multiple layers of management, administration and academic roles. Being owned by a corporate organisation was described as providing for a more "*effective management system*" with decision making within a committee structure promoting "*consensus*". Mike described the owners as having "*put a lot of resources*" into current educational best practice and described how they saw themselves placed against other institutions.

> we are more concerned with our outcomes[11] than some of the non-profit, public sector, because that is what students look at. It is consumer driven. (Mike)

This articulation was demonstrating the new expectation of the student as 'consumer' and illustrated the financial importance for the organisation. In this discourse students were seen as both raw material and product. Mike explained the situation further.

> accreditors across the nation of which there are 100's, you think about all the different professional associations they have not caught up with programme review, they are still input oriented rather than outcome oriented.

No longer inputs, it was the HE requirement for outcomes as a standardising procedure to direct programme reviews that was being embraced by the institution proactively. This institution had also incorporated competencies from mainstream medical education into their course.

> So the revisions we have made to our clinical examinations has been based on the merging of the ACGME guidelines, the American College of Graduate Medical Education, they have clear competencies on what health care providers should know and do for health care in general which we use. And also the Institute of Medicine guidelines - a lot of their outcomes. (Mike)

This match to the ACGME competencies would provide evidence of best practice and align the institution with other health care providers. This is the same impetus that I discussed in the previous chapter that could be seen in the

---

[11] Outcomes here do not refer to learning outcomes but more to institutional results and statistics.

changes to the updated ACAOM (2019) documents. As a political and economic strategy it would give the institution a competitive edge and enhance their status as a progressive educational institution.

In another institution the need to meet statutory regulations was manifest in the format adopted in their assessment instrument. The institution had two instruments in circulation that listed an identical number of competencies except that the shorter form did not break them down into multiple learning outcomes. It was the longer instrument that was to be applied by the supervisor observing a student treating just one case. In untangling the difference in the two documents it transpired that the 'long form' had been written for a very different practitioner group, the result of the "*introduction of the national registration system*" and subsequent "*grand-parenting*" procedures for older or overseas practitioners. Why an assessment for registering trained practitioners based on one case which was not compatible with that for assessing students that needed to be based on observation over time, was being used had to be questioned.

> *In terms of using the form in clinic it is a really <u>long</u> form and it is quite complex but it is consistent with those used in other <u>health professions</u> looking at skill based assessment processes. (Maude)*

Maude was aware of the forms' origins and length, complexity and inconsistencies, but its alignment with other "*health professions*" and the requirements of statutory regulation seemed to be the key to its development. The result of this, however, was that this instrument with its multiple learning outcomes, was not being enacted or could not be enacted in the way instructed. The supervisors' involvement (or lack of involvement) in this new WBA will be returned to in the next chapter.

The drive for status was also an impetus behind the move for the programme in the UK to enter into a validation process with a university and thus open themselves up to a more intensive academic scrutiny and greater standardization. This further validation process was consumer driven - i.e. students wanted the higher status of a degree qualification and it gave students access to grants and thus conferred a competitive edge in recruiting students.

The impact of the particular managerial and pedagogical theories emanating from the external stakeholders of the different countries, either accreditors or validators or corporate owners, was being dealt with in different ways by the institutions. What was in common was that the assessment instruments and the language incorporated had a direct link to these requirements, on the students' experience, and on WBA.

5.7 The Marginalized Discipline

If the successful use of assessment instruments resides in the assessors (not withstanding the critique I presented in the literature review) then how an institution recruits and trains its supervisors will impact on that success. While all institutions had requirements that clinical supervisors needed to have been in practice for 5-7 years and be registered practitioners of CM, no institution asked for specific teaching qualifications and there was very little evidence of any training in clinical supervision or the use of assessment instruments, especially in two of the institutions. Managers commented that:

> *Everyone has the [discipline] knowledge but they can't teach. (Mary)*

> *There is always the assumption, and this is the problem with a lot of Universities, that if you have a qualification in something then it means you might be efficient in teaching and learning. And it is not always the case. And the assumption is because you are a registered Chinese medicine practitioner that you have a minimum 5 years experience then you can also be a supervisor. So obviously the emphasis has been on disciplinary skills than supervision skills. (Maude)*

While Mary and Maude were aware of the implications of a lack of teaching skills, any training for supervisors was "*very basic, a taxonomy",* an overview of the systems and documents, or "*shadowing"* for a day. In relation to assessment, Mark reported:

> *I think … that the clinicians need much more training in what it means to be an assessor. I think that part is very under emphasised in their own minds. I think they see themselves more as a <u>mentor</u>. I want you to learn the things that I know about how to be a good clinician.*

The linking of the role of assessor to the identity of the supervisor as a mentor illustrated further the awareness of the difference in these roles. Mentoring can be seen as a valid pedagogical strategy for WBL (Higgs, 2012a) as it can bring with it the more tacit and embodied traditions and narratives and ethos of the discipline. Eraut (2010) recommends shadowing as important for learning what underpins systems and routines and also how intuitive decisions are made. However, as methods to embed the new recruit into the community of practice, they were not being used in any strategic way.

One manager, Mary, who recruited the supervisors, was not interested in the applicant's teaching qualifications or of their specific knowledge of CM but their "*ability to tell a story*, "*get information across*", to "*organise*" and "*run the clinic shift*". For Mary all applicants had the knowledge and skills of CM but it was communication and organisational skills that were key for being a supervisor. Knowledge of learning outcomes and the ability to make judgements in assessment decisions were not proposed as important. The tension between being CM-focused or education-focused, between being focused on the treatment of patients as opposed to the assessment of students as well as being motivated and dedicated is an area that will be developed in the next chapter as the supervisors' story unfolds.

5.8 Contracts and Communication

The impact of supervisors' work-contracts on communications between management and supervisors and across supervisors, introduced a further factor impacting on WBA in all the institutions. For one institution, in line with their accreditor, there was a conscious choice by the managers to have a range of practitioners in clinic, the rationale being that this would provide the diversity of approaches to practice students needed. As discussed, the employment of multiple supervisors that this system created allowed for the input of marks from multiple supervisors. Further, short and small contracts, it was explained, protected the institution against the problems incurred if a supervisor left and needed to be replaced. While faculty meetings took place three times per year the time spent on discussion of clinical issues was said to be very brief. It was

recognised that the limited employment contract might be impacting on the level of thinking around pedagogy.

> *They are all part-time, they are not <u>engaged</u>, they don't understand the level of thinking … and their biggest shortcoming is lack of depth and knowledge about the college and teaching the next generation. But they don't have the <u>time</u>, have the understanding of HE and changes in East Asian medicine and health care (Mike).*

A lack of time was presented by Mike as a factor in the supervisors' lack of understanding of institutional processes or pedagogy. That, combined with a lack of training in education, would limit the development of skills of assessment. Any knowledge of assessment, therefore, came primarily from the supervisors' experience as students with it being considered particularly useful if the supervisor was a graduate of the institution as then they would be "*on the same page*". Remember, 50% of the supervisors were trained in their own institution. However, Mavis, another manager in the same institution, contradicted the issue of lack of engagement giving another slant on the position of the contract.

> *They earn a lot more money in private practice, they want to be connected to a <u>community</u> and give back, they are <u>not</u> <u>disgruntled</u>. I find that even a member who is only here for 3 hours per week, I can get them engaged because they are <u>motivated</u> and <u>dedicated</u> to the school. (Mavis)*

The supervisors' engagement with their role was related to the sense of "*community*" provided by the institution and "*not disgruntled*" referred to the relatively poor pay. Within all institutions the supervisors were part-time and had busy private practices. While Mike was concerned with the supervisors' lack of teaching experience and thus engagement, Mavis considered that the supervisors' role as practitioners within the community of CM was what supported their engagement. This opinion resonates with the research of Buchel and Edwards (2005) who report clinical competence and enthusiasm for their educational responsibilities as the key attributes of effective clinical teachers.

In two institutions the supervisors were on part-time tenure, however, while tenure did seem to allow for better communication amongst supervisors it did

not necessarily mean that there was adequate time allocated to meetings that concerned curricular issues. With their time focused on clinical teaching the institution that had introduced a new assessment instrument acknowledged the problem.

> *They* [the assessment tools] *were written about 2-3 years ago and they have kind of been phased in. This year we have tried to use them. The difficulty is that the clinic is so busy it is hard when you have got something as detailed as that to do justice to the form. It is a work in progress let me say that. (Mark)*

This statement emphasised how the busyness of the authentic work setting was impacting on whether supervisors could assimilate the new assessment instrument, and explained why there was still confusion as to its very structure. The position taken by Mark, however, was in relation to the supervisor's incorporation of the instrument rather than any concern about the construct of the instrument. As highlighted earlier, the amount of time the supervisors spent in clinic supporting students treating patients, was seen as evidence of students' learning. Mark continued this theme of busyness.

> *The main thing is time as the clinic is so busy. Do they have the time to sit down and do this* [assessment] *formally or do they do it quickly. They do have a good idea* [of their students] *if you talk to them and I have several times, how is that person going. They know, they do know as they are in here all the time across lots and lots of clinics. (Mark)*

Quality time spent with students, that is getting to know them and giving feedback, was seen as more valuable than supervisors having time to administer the new assessment instrument. The statement could also be interpreted as time with students in practice allowed for a better global assessment of students than their evaluation of them against the learning outcomes. I will return to this argument in the next chapter.

Even if on tenure supervisors were often in on different days and Matt voiced an awareness of needing to calculate in meeting time.

> *So what we came out with from that was the need for them to have some other time that's not geared towards particular assessment or finishing off the year, or anything like that, but that is separate development time, for them as a team, so that they can share best practice. (Matt)*

Time to meet and share good practice was seen by this manager as essential a component as the discussion of the paperwork and determining students' marks. O'Neal (2016) sums up the importance of communication.

> The clarity of standards established to grade authentic assessments do not appear to be nearly as important as the discussion of those standards that happens when assessors interconnect (p. 66).

Summary

The answer to my question - how is CBME being enacted in the institutions (documents and managers) and what issues does this present? - was complex. The documentation of all institutions promoted CBME in the application of a psychometric grade to denote achievement of students' performances aligned with published, standardized and atomized learning outcomes. Leaning more towards the requirements of other health professions, especially in the USA and Australia, the learning outcomes lacked recognition of the often subjective, interpretive and inconsistent nature of professional work and further were written in a language that stripped out the essence of CM. The format of the assessment instruments with their multiple outcomes, and for two institutions a lack of criteria to differentiate standards expected, promoted grade inflation and failure-to-fail. This constrained the validity and reliability of the psychometric grade even while 'fairness' was being enacted in different ways. A recognition of the importance of an authentic environment to carry out performance-based assessment (Rethans *et al.*, 2002), was manifest in the extensive investment in their teaching clinics both in terms of the buildings and resources and in the supervisors.

While the institutional documents portrayed a leaning towards standardization, the managers were aware of a more complex story of WBA. The lack of value allocated to the WBA as shown by grade-inflation and failure-to-fail, was more to do with the subjectivity and inconsistency of working with students and patients than a devaluing of practice experience The educational training of the supervisors (or lack of it), induction (or not), involvement in the development of differentiating criteria (or not), and moderation (or not) were factors enabling or constraining a valid and reliable implementation of CBME. However, the

supervisors' credentials in CM and the hours supervisors spent with students supported the authenticity of the environment as a major factor in the enactment of WBA. Behind the lack of objectivity of the marks being awarded lay the nature of supervisors' contracts and the busyness of the clinics. These issues will be pursued further in the next chapter as I turn my attention to the supervisors' experience and how they understood the place of measurement, of holism and the role of CM in their enactment of WBA.

Chapter Six: Supervisors – personalized competencies

> … it will be the talents, interests and capacities of educators who will be pivotal in determining whether the kinds of outcomes sponsors of educational institutions seek are achieved, the demands of external bodies are realised and students' needs are adequately understood and responded to in ways that are directed to their interests … (Kennedy *et al.*, 2015, p. 5)

While professional bodies articulated the competencies that they expected graduates to meet, and managers produced learning outcomes and psychometrically constructed assessment instruments, the role of the supervisor was crucial for interpreting and manifesting these abstracted and de-personalised statements in their direction and assessment of students in direct contact with patients. In this chapter my interest turns to how the standardized format of the assessment instruments was being understood and implemented by the supervisors. What were the issues around the interpretation of the learning outcomes and objectivity and judgement? How did the supervisors blend their roles as assessors of students' competence and practitioners of CM in an authentic environment? In this chapter I will consider the story being told about WBA from the perspective of the supervisors.

6.1 Personalized Competencies

The variability and inconsistency in assessors' use of assessment instruments in Universities has been documented in a study by Bloxham *et al.* (2015) who identified five reasons that might account for this variation namely: the use of different criteria to those published; different understandings of the criteria; different understandings of standards; diverse sub-criteria and different weight and value being given to the criteria. This study supported those findings and considered further aspects beyond the criteria to the format of the instrument and how assessors understood the assessment process.

While the supervisors articulated their expectations of students quite clearly only two of the thirteen made a direct reference to the published learning

outcomes. Some of the competencies supervisors described as key expectations of students such as "*safety*", *"knowledge and skills", "critical thinking", "diagnostic acumen",* and *"good communication skills"*, were contained in the published learning outcomes. Their awareness, however, of these competencies seemed to emerge from their personal view of the practice of CM as opposed to any sense of the institution's official learning outcomes. Rather than describing this as a lack of accountability as proposed by one of the managers, my interpretation was that they were acting on 'personalised' competencies. The supervisors presented a different view of what competent practice might look like as illustrated in the following extracts.

> *The student being able to have <u>insight</u> into that beautiful wonderful human being who is in front of them who they are trying to treat, that is what I look for, that is what quality is … how do you measure <u>devotion</u>, how do you measure attentiveness, things like this? (Chris)*

> *But once they get that click like they get <u>enlightenment</u> and then they are good. (Frank)*

> *As long as they are present, they are appropriately attired, they are <u>respectful</u> of the patients, they have performed to the best of their ability to the level of internship they are in. (Harry)*

> *They need to have <u>a loving and caring heart</u>. Some people, students, I know they have good skills but they do not really look after the patients very well… and in the future I can tell they will probably focus on the money not on the patient. But most they need to be <u>willing to help others</u>. This is very important. And the other thing they need knowledge and skills and experience. They learn from clinic and need to help others. These are the things I need to know whether they have this or not. (Julie)*

> *Honesty, <u>dedication</u> to what you are learning. (Ned)*

> *I like someone <u>diligent</u>, hardworking, hands on, willing to learn and issue questions, inquisitive in learning. (Kate)*

> *There are some people who become <u>over-confident</u>, they think because they've finished the course, they know everything and rather better than everybody, all the tutors combined… they are the most dangerous I think. (Greg)*

Nowhere in the published learning outcomes was there reference to "*insight*", "*devotion*", "*dedication*", "*enlightenment*" or a "*loving and caring heart*", not even negatively to "*over-confident*". The extracts above showed that supervisors

were relating to a view of the student in relationship to the patient and to their own beliefs and experience of practice, a judgement, perhaps equally as valid as, but different from, the published outcomes. These statements imply that the different values expressed by Ho *et al*, (2011, 2012, 2014), issues of integrity, relationalism and social and personal roles, were finding expression. Research conducted by MacLeod (2011) found that medical educators and medical students described the need to develop two competing identities in their training: that of competent doctor in which the approach was competence as knowledge of biomedicine and clinical skills and the caring doctor emphasizing social issues. MacLeod describes knowledge as taking precedence over caring in the WBA carried out. In this study caring was a major aspect of the supervisors' assessment of competence even though it was not articulated in the learning outcomes.

The extracts above encompass some of the aims written into the institutions mission statements and general descriptions of the clinical environment as discussed in the previous section that had not been translated into learning outcomes. Hodge, Mavin and Kearns (2019) concede that everything relevant to practice cannot be captured in competency statements and that students, within their immersion in the practice, will gain these aspects. Their concern, however, is that those things thus omitted are de-valued. Brown's (2000) findings in nurse education reported that 'the assessment of student performance in clinical practice is not restricted by pre-determined behavioural learning outcomes' (p407). Likewise, the supervisors in this study were using the more tacit aspects of practice in the determination of grades, their value thus acknowledged as shown in the following statement.

> *Assessment is supposed to assess achievement, not <u>effort</u>, not the process, not making the <u>effort</u>, but when the student is <u>making an effort</u>, I give them one or two marks anyway. (Kate)*

For Chris it was the learning process that guided his view of students

> *… I don't want to say anything negative, they do not do anything egregious, they really don't. They don't do anything outside of the <u>normal process of learning</u> to be in clinic. (Chris)*

"Making an effort" along with the "normal process of learning" turned the attention of assessment to the learning process or journey not the product. Almost all supervisors (and managers) described the students as wanting to be in clinic and wanting to learn and their role was to guide them. Diligence and effort, however, were also not factors in the published learning outcomes.

Personalized competencies being applied to assessment had other manifestations. Some supervisors saw learning as not just applicable to the time in clinic but projected forward beyond their time as students.

> So if the person is keen to learn, they have already ticked my box. And if their theory is a little bit weak, I feel it doesn't really matter, you can learn, over the year, after graduation. If theory is weaker, you can pick up later. (Kate)

> I do not like grade inflation but I do not see any reason not to give this person – you passed and did fine. But is it a fine B or a fine A. It's hard to say as you are in the process of building skills. It's going to be rocky and it will continue to be rocky once they are out there. (Chris)

That students were seen as early in their development as practitioners of CM and would continue to learn after they had graduated, impacted on what these supervisors considered crucial to their assessment. They were giving different weight and value to different criteria and had a view of competence as 'good enough' (Yorke, 2005). That learning would extend into the future, what Eraut (2004) describes as a time trajectory of competence, changed the supervisors' discrimination, modifying the meaning of competence as knowledge and skills. The lack of agreed criteria amongst supervisors would certainly limit, however, any possible inter-rater reliability.

Two supervisors were aware of the assessment instrument as an institutional requirement influenced by outside stakeholders.

> they have academicians who find these tools in academia and pull them over. Because they want to and it is understandable that they want to feel part of the medical community so they are using evaluations that are not Chinese medicine, based on Chinese medicine principles. They are based on normal milestones in the world of clinical practice. Perfectly understandable right? (Chris)

> This is a new one [new assessment instrument]. Before we only use this one [short form] but later on because to meet the board requirements or

> *something like that they develop this form* [long form]. *This is in more detail. (Julie)*

These quotes display the supervisors' awareness, if not support, of the priority of the institution in their dealings with external stakeholders. At the same time, however, the supervisors had adapted their own assessment process to militate against the difficulties that ensued from the construct imposed and one not based on CM principles. Julie justified applying a different time construct to the assessment process.

> *It* [the assessment] *is supposed to be at only that time they are doing the consultation for that single patient but I think we cannot just according to one treatment to give the marks we need to look at their general practice. Because some students are very nervous when they are treating the patient and are nervous and are making some mistakes so I still look at their general behaviour in their practice to give them the marks. (Julie)*

This statement ties in with the analysis in the previous chapter about the assessment instrument produced by the managers for a different group of practitioners. The instructions were misleading and conflicting and so the supervisor was managing the restraints imposed by the assessment instrument in order to be more fair on the student. While Govaerts, van der Vleuten and Holmboe (2019) are concerned that standardization might create a 'lack of ownership and agency' (p70) on the part of the assessors, this sitting between both experiences, the acceptance of the institutions' role and their own, manifests what Cantillon, Dornan and De Grave (2019) describe as a 'mutuality' between external and internal values thus maintaining their own personal agency. Supervisors could be seen to be using their discretion as to how they assessed students' competence.

While applying personalized competencies, supervisors were also assessing students against an holistic understanding of practice.

> *I look at them as a whole. The knowledge and also how they interact with the patient because it is very important because it does not matter how wonderful how much knowledge you have how much skill you have if you don't communicate with the person the patient does not follow so they don't have a same result. (Frank)*

> *I see them together; I do not go through each one* [learning outcome]. *Just overall. These students follow us for a whole semester, or a whole*

> *year. We know them very well. Sometimes I don't need this one* [the assessment instrument] *I know how to do that. Not in detail. It is hard to go through each single thing. Unless they give us time the whole day I am sitting here doing this maybe I can do in detail. (Julie)*

What these extracts showed was that the supervisors were applying an holistic judgement, generating a mark for the student's overall clinical performance over time, and not broken down into atomized learning outcomes. This was the general approach to assessment for all supervisors irrespective of the assessment instrument. They described 'knowing' their students, an embodied knowledge developed over time working closely with students. Supervisors were applying personalized competencies within an holistic assessment process.

That supervisors applied a holistic assessment of students was not just the result of their attitude to the atomized learning outcomes but could be seen as the result of the construct of the assessment instrument.

> *And the second hardest, I think, is the balance of… some people can be really natural, therapeutic behaviour, very sympathetic, empathetic, very professional demeanour, but clinical, analytical process is slightly weak, theory side a little bit weak, how to balance these few, because in our marking criteria there is no percentage - this is 25%, that's 25% - we didn't give that. (Kate)*

What Kate was struggling with was how to place a numerical value on the different criteria. The assessment instrument did not apply a percentage value to each and, as described previously, there were no instructions around compensation. I propose that this might relate to the view of Ginsburg *et al.* (2010) who challenge the linear format in which instruments are written that present each competency as equal. It is this 'visual rhetoric' that they consider is 'inconsistent with the way faculty actually conceptualize and express their opinions about the performance of their residents' (p.785). Thus the instrument itself generated the holistic judgement in juxtaposition to its linear and atomized format.

A study by Bloxham, Boyd and Orr (2011) concluded that there is a contradiction between the stated policies and actual practices in HE in relation

to criterion referenced assessment and marking. They propose that this is not a 'cynical rejection of accountability or a determined adherence to conventional practices' (p667) or a dereliction of standards (Knight, 2007). It is better understood as the tension between manifestations of accountability as shown in the published assessment tool and professional judgement as manifest by the supervisors.

6.2 Objectivity

The aim of a psychometric format of assessment instruments is to support fairness of marking across multiple assessors. However it was the very construct of the assessment instrument that drew criticism from some of the supervisors especially in relation to the psychometric grading and atomization.

> *The latest one, the sophistication of the questions was really good* [but] *there were far too many of them but far, far, far, too many and the actual numerology in grading them was a complete improvisation, it was just beyond belief… So it made it very difficult to want to take seriously these questions that were asked. But there were just too many questions and just too much detail. It would be impossible for me to truly answer them with any sort of veracity. (Chris)*

> *It has decimal systems that add up and divide into this number is that grade and that number is this grade. And the final assessment - you are just meant to give them an 'A'. (Lily)*

> *I think it's a more complicated version, but for me, it's actually too complicated, there are too many words. (Greg)*

As discussed in the previous section, some of the assessment instruments had very complex marking systems and all had multiple atomized learning outcomes and it was this aspect that lead to the supervisors under-valuing the instrument in terms of recording achievement. The lack of value emerged from its complicated construct of psychometrics and atomization rather than it being the supervisor's lack of skill in assessment as described by the managers.

Another approach to the instrument was voiced by Bee as:

> *Assessing is easy it is just the paperwork that is cumbersome but now it is online and usually pretty quick. They have those tick-off the boxes.*

This statement presents the tick-box format of the instrument as easy to use. However, when I asked Bee if she went through the assessment form with the student her reply was "*you do it on line now*". This supports the argument that the tick-box format of assessment instruments, especially on-line, is what devalues the instrument and removes it from being a source of learning for the student.

The supervisors supported the dilemma around the subjectivity of WBA raised by the managers. The main aspect of student performance that would lead to a lower than 'A' grade was stated as lack of attendance or not filling in the paperwork – both objective measures. The issue of objectivity was also illustrated by the difference perceived between clinical assessments and written assessments.

> *This* [WBA] *is in relation to their clinical practice and their general behaviour and their skills. This one I feel is <u>all in my mind</u> but that one* [written case study] *is on paper. I look at the paper. The case study is complicated and makes them to think but according to the answer.* [In] *the case study there is an answer as well I have already done and I see if they match my answer or not. But this* [WBA] *is <u>more flexible more real</u>. (Julie)*

Julie was describing the difference between the marking of a pre-written case study and assessing a student in the moment of the clinical experience. In the written assessment the student had committed their answer to paper and so there was something objective to read and the supervisor could mark against their own 'model' answer. The clinical situation was uncontrolled, inconsistent and demanded more interpretation. This explanation brought into focus the lack of objectivity possible in WBA that is in the moment, influenced by the authenticity of the context.

6.3 Failure-to-fail

All the supervisors except one reported that they had not failed a student in their WBA. While practising was seen as key to students' learning it was the subjectiveness of the WBA that supervisors described as a reason for 'failure-to-fail'.

> *I need to fail this student, because the supervisors' feedback is that they are worried. He [the student] said where's the evidence? There was nothing… so I couldn't fail him. (Diane)*

The lack of written evidence as a means to objectivity impacting on a fail decision voiced by Diane supported the managers' concerns as discussed previously. In relation to WBA, research in the health professions has shown that it is extremely rare for students to be failed (Hughes, Mitchell and Johnston, 2016; Hunt *et al.*, 2012) due to the possible personal or professional impact (Yepes-Rios *et al.*, 2016) and also fear of litigation (Gingerich, Regehr and Eva, 2011; Hawe, 2003; Yorke, 2005).

As discussed in the previous chapter, in the UK institution a moderation meeting was used to make final decisions on grades and as a means to deal with the subjectivity and biases that could ensue from WBA. Dudek, Trier and Hauer's discussion (reported by McColl (2017)) around the issue of 'failure-to-fail' consider the sharing of responsibilities for making decisions on student achievements a key step in the use of data from CBME assessment instruments. Moderation can also be seen as a critical phase in an institution's merging of their own philosophical and individual mission with that of the accreditor and of a merging of ideas between assessors. In the other two institutions the supervisors had not been involved with the writing of the learning outcomes or the assessment instruments and their vocal critique of the tools as "*unworkable*", "*too much hassle*", "*too long*" was powerful.

The following extract illustrated a similar rationale as that the managers voiced as to why students were not failed.

> *The reason why it is rare is that the students spend a lot of <u>time</u> on the programme, it is a lot of <u>money</u>, a lot of <u>effort</u>, so you get students who for the most part are very <u>serious</u> and for the most part they have been <u>vetted out</u> before their final year. Somewhere in the observation or assistant level. (Ellie)*

There are three features in this extract. Firstly, was that weak students would have been failed before this stage of clinical training in assessments that could be more objective. Secondly, manifesting the 'input' curricular model, learning could be assumed because of the extensive "time" students spent in clinic.

Thirdly the cost and effort students had to outlay would discourage the less serious student. These beliefs, articulated by others as well, were impacting on how the supervisors viewed students, assessed them, and (almost) always passed them. As challenge to this, two supervisors did voice concerns about weak students.

> *I don't know how it works here. It seems to me that there are students who have got to this place and in my personal opinion should never have reached this point and shouldn't go on from here. It seems that that is more theoretical than practised … And I don't know, I have not been here long enough, I don't know if it is my responsibility to fail them or another thing that happens in their academic records. (Lily)*

> *And others that I will need to say oh my god I will need to call you a colleague someday and you will actually do what I think is a disservice to the field. But they have enough gumption to get the assessment points and get through. (Mitch)*

While Lily challenged whether students had reached an appropriate level of competence before the clinical component of their programme, Mitch raised a concern that linked to the construct of the assessment instrument and its inability to discriminate between competencies. As described above, and supported by Yorke's (2011a) analysis of WBA, the assessment instruments did not disallow compensation across competencies and so lead to students being passed with individual competencies at fail. But the quotes also illustrated a lack of awareness as to where the responsibility for failure lay. In two of the institutions final WBA grades were collated by a third party and what these supervisors displayed was a lack of awareness of the impact of their assessment on the final results. Thus it was the bureaucratic process that also impacted on the reliability of assessment.

Supervisors did provide quite detailed descriptions of how they assessed students but this was not as Knight (2007) describes as assessment for 'warranty' but more in line with assessment-as-feedback. This is what Boud (2007) proposes as the original idea of feedback which needs to influence the nature of the output in-the-moment, that is, 'feedback-for-learning'. The marks, therefore, that supervisors in two of the institutions were giving reflected a general competence, not a differentiation between students (except in the

smaller institution which I will discuss further). According to Lynch (2014) many medical schools in North America are abandoning grading with letter or number and beginning to use pass/fail models of assessment combined with a descriptive and dialogic format in order to meet the problems of applying objectivity in WBA.

6.4 Assessment-for-Learning

Supervisors were immersed in the treatment of patients, watching the performance of the students and giving immediate and appropriate feedback verbally in relation to each patient, ensuring effective and safe practice and maintaining a thriving teaching clinic. The supervisors described the process.

> *I will come in and feel the pulse and look at the tongue and will have the students feel the pulse and explain what they feel and I will explain what I feel and then we will come up with the diagnosis on the spot and the treatment plan. So I am in the room for most of that. (Harry)*

> *I mean the formative assessment kind of goes on every week they appear in clinic. And with every patient that they see, they are getting feedback, and you are both pushing to <u>develop their thinking</u> about the case, and sometimes different aspects and so on. (Greg)*

> *Even though we have these formal assessments, assessment is <u>on-going every day</u>. (Diane)*

> *But the student follow me for whole year and I know them very well. So I tell my student this assessment <u>the paper</u> just a procedure. I assess you every time. But I go with them, they ask me questions and I really know how they improved or not improved. (Ann)*

For the supervisors, feedback was "*on-going every day*" guiding the students, helping them develop confidence, their thinking and their technical and diagnostic ability. A measurement of the students' learning in relation to the assessment instrument, "*the paper*", the product of learning, was not the supervisor's focus. Nor was assessment to differentiate one student from another as shown in the following extract:

> *their relationship with the patients is what is important, not who is doing better so how could you possible measure that? (Harry)*

119

Given the difficulty of coming to valid and reliable psychometric decisions on WBA due to contextual and relational influences, my analysis showed the supervisors bypassing this. They saw their role as carrying out 'assessment-for-learning' as a means to inform students of their strengths and weaknesses, as opposed to 'assessment-of-learning'. In so doing they aligned with much of the literature that recommends that this should be the focus of WBA (Boud and Molloy, 2013; Sambell, McDowell and Montgomery, 2012; Schuwirth and Van der Vleuten, 2011b).

The analysis I have presented above raised further questions, however, of what might be mediating the relationship between the supervisor and the assessment instrument and process that was provoking such a de-valuing of it. A number of possible factors came to light that related to the personal relationships formed between supervisors and students, to employment, space, time and identity.

6.5 'Knowing' students

As discussed in the literature review, WBA is a subjective process involving supervisors making judgements about a student's performance and this can create a tension with the standardization and psychometrics of any assessment instrument. The considerable time supervisors were spending with students within the authentic clinic environment allowed for a high level of intimacy to develop.

The key skills supervisors described as being essential for their role included knowledge of Chinese medicine including technical skills, a "*good grasp of the basics*" and "*superlative diagnostic skills*''. This latter aspect was described as having a "*practitioners eye*" and involved the ability through observation of the patient and skill at pulse taking and palpation, to very quickly see if the student was making a sound diagnosis. Supervisors also described appropriate pedagogical approaches such as the need to "*guide*" the student, to give "*positive feedback*" and "*not criticise or constantly correct*". "*Helping students make connections*", "*awareness of cultural differences*" and being "*sensitive of the students learning style*" were also presented as being important. Four

supervisors described needing "*confidence*" or "*lack of anxiety*" and three that managing "*time*" was important. Behind these skills other factors were voiced that mediated the relationship with the student, namely 'loving' their job and 'knowing' students.

> *There has to be a significant <u>comfort</u> level with your own skills as a practitioner with your own grasp of the medicine and much more important than that is a <u>love</u> of teaching and a <u>love</u> of patients and a <u>love</u> of medicine; if you don't have the love for all three I would find it very easy to burn out. (Lily)*

> *First of all knowledge, your professional level and second your <u>love</u> [of] students and your love [of] teaching and you are happy to pull out your knowledge; and <u>knowing</u> your students. (Ann)*

> *You have to <u>know</u> your student. (Ian)*

To be highly skilled and confident was linked, not so much to the skills of supervision but to a "*love*" of teaching and patients and the medicine and the students; it was this love that sustained their role as a supervisor. They also described 'knowing' their students. Supporting Dornan *et al's.* (2005) research, the interpersonal dynamic of the learning relationship between the supervisor and the student was a factor that almost all supervisors drew attention to. Broudy (1972) describes the relationship between student and teacher as 'philetics', a loving concern for the students' development both intellectually and as an individual suggesting that 'the love relation removes the psychological blocks to learning' (p.253). Frank described this mediating factor in a very Chinese philosophical way.

> *but the clinic is very interesting as they are the one choose you according to their time and preference and then their energy is more <u>harmonised</u>. It is very interesting that kind of dynamic because they choose you and they have, their <u>energy</u> already there, because if their <u>energy</u> is good then the outcome seems better because they are <u>willing</u> to listen and I feel I am able to really discuss what I think. Sometimes students are coming and their energy not good and then they don't really listen or do not understand what you mean; it is not that they are not smart its just they don't understand and we don't communicate.*

Firstly, this extract described the way Frank perceived the students' choice of him to be their supervisor as one of "*willingness*" that would therefore lead to a "*harmonious*" rapport and so better learning. Secondly it highlighted that the

relationship between Frank and the student had an energetic, an embodied dynamic. This was not an imparting of knowledge and skills from supervisor to student but a reciprocal relationship described in terms of "*energy*", a concept at the heart of the philosophy of Chinese medicine. The intimacy of the relationship the supervisors were relaying was also captured in expressions such as knowing about their "*family life*'" their "*financial pressures*", knowing about the "*debt*" the student would be going out with after training. This knowledge of students was made possible by the sustained period of time supervisors spent with each student in the authentic setting, "*working as a team*".

This intimacy of "*knowing your students*" had a direct impact on the assessments made. Trede and Smith's (2014) research into WBA with physiotherapy supervisors also concluded that the knowledge gained from working with students over a period of time led to 'fairer' assessments. While this might be the case for the student, the intimate relationship built up presented a dilemma for some supervisors.

> *If we assess our student we be a little bit embarrassed. You understand. We love the student and more likely to give the student more higher mark. And I wish we have other lecturer coming down and do* [assess] *independent* [ly]. *But for the students they are with us for one year and we have personal, we build up personal relationship, we love the student and student love us. So we be a little bit embarrass. We tick it more here* [at high end of criteria]. *(Ann)*

What Ann described so clearly was that the mark she gave was seen as a validation of the relationship between the supervisor and the student with a low mark construed as a betrayal of that relationship, a judgement of personal worth (Hawe, 2003). It was not the teaching and learning that supervisors had problems with, nor the fairness of the marking, but the turning of their knowledge about the student's performance into a numerical grade and the impact of that on the relationship. Supervisors were also concerned they might undermine students by putting what could be viewed as negative feedback in writing.

> *Sometimes we just tell them, I don't want to mark here as they* [the student] *feel bad. And when they look here they feel bad so I just leave them and not put a mark here. (Julie)*

A psychometric approach was seen to have a detrimental impact on the teacher-student relationship and so Julie avoided any actual marks. In order to minimize the impact on the relationship supervisors could be seen to commit little to writing preferring to give oral feedback and for the final assessment gave a global assessment allocated at the higher end of the grading criteria. Hodges (2012) considers that feedback can be limited when the psychometric discourse is dominant, and recommends a 'narrative' approach especially for longitudinal assessments in the workplace. My analysis of these institutions suggested that, while any written narrative was limited, extensive oral feedback was provided in relation to each patient.

The reciprocal relationship between the supervisor and the student was manifest further.

> *I have been fortunate that I have not ever had any very difficult students to work with over these many years and many clinic shifts. I really feel I have been <u>lucky</u>, because I have heard stories about other students. I have had no reason to fail anyone or even give someone a poor grade because they <u>like clinic</u> they want to be part of that experience. (Chris)*

> *And my shift there is actually a wait list for my shift so I tend to get very good students, <u>all stars</u>, students who want to perform and want to work well want my shift because they know I will want them to do that. The reward they get is carte blanche – they can do what they want to within certain limits. (Mitch)*

This expression of being "*lucky*" with the students they had worked with, that they were "*all stars*" was mentioned by other supervisors as well further implying something beyond a formal teacher-student relationship but one where a mutual rapport allowed for an harmonious dynamic in the clinic. This feeling towards the students, as illustrated by Chris, manifested on the actual grade they might give. As Ginsburg *et al's*. (2010) research shows, the social encounter within a clinical workplace contributes to the unevenness of assessment with the emotional reactions of supervisors and students adding to the inherent subjectivity. In relation to something as complex as clinical practice, it is problematic to think that competence can be viewed as a 'stable trait' and that a 'single true score' can be established by the supervisor as a dispassionate observer.

The findings above support the report of Kilminster and Jolly (2000) that in terms of the effectiveness of supervision, the relationships between the supervisor and the student are even more important than the supervisory method used. Each supervisor had their own approach to supervision mostly gained from their own experience of being students themselves. Shanahan, van der Vleuten and Schuwirth (2019) propose that it is around assessment that the supervisor-student relationship often breaks down. Preston (2017) and Hyland (2017) have warned that the requirement for evidence of achievement changes the possible relationship between the supervisor and the student from the process of learning to an objective measurement. These propositions were not evident in this research as achievement through an objective assessment was not a priority for the supervisors. By working intimately with students in direct relationship with patients and providing assessment-for-learning in regular verbal day-to-day feedback, it was the supervisors' 'knowing' of their students that was being converted to a grade, a high grade.

6.6 Employment and Time

The managers presented one of the weaknesses of assessment as being that the supervisors were not enough aware of institutional procedures and needed much more training in the role of assessor. What manifested through the interviews with the supervisors was more complex.

Each institution had different employment systems and it was these differences that helped illuminate the process of assessment. In one institution supervisors were employed as sessional workers for anything from 1- 3 shifts in any week with no fixed contract beyond one trimester. This was a conscious decision on the part of the institution as described in the previous chapter. As well, part of the institutions' quality assurance system was for students to grade the teachers, and one of the managers had described how being *"unpopular"* might lead to a supervisor no longer teaching. This was also raised by one of the supervisors.

> *And I think part of the issue is that our jobs are at stake as we have to be requested as supervisors, as much as we give students feedback the students give feedback on us. (Lily)*

What Lily was describing was the power students had over supervisors. If supervisors gave low grades then students would in turn give a poor evaluation of the supervisor and by not choosing their clinic shift, that supervisor would not be employed. That failing students might negatively impact their own tenure was cited by Yepes-Rios *et al.* (2016) in their systematic review of health profession education and one of the reasons behind 'failure-to-fail'. Lynch (2014) has described how grading is increasingly being used not just to assess students, but academic staff as well, linking this to the neoliberal desire for accountability through measurement impacting on the conduct of teachers, and raising of grades.

The sessional nature of the supervisor's employment along with the busyness of the clinic day limited the time available for meetings with management and with each other to discuss the process of teaching and assessment. Most supervisors reported that they did not talk to other supervisors – "*there is really no time outside of a cursory hello" (Harry)* – and left the institution immediately after their shift.

> *No we are running around like headless chickens, we are all very busy. It is always a mad dash. We don't have that <u>sense of community</u> in the clinic as we had before…*
> *But it's <u>busy</u>. In the previous set up the <u>faculty lounge</u> was on the same floor as the clinic so you could go to the faculty lounge and put your feet up and there would be another person there. (Chris)*

This extract described the lack of time and the busyness of the clinic, as preventing the "*sense of community"* that allowed for a sharing of experiences. It also introduced the architecture of the institution as having an impact on relationships. It was the new location that separated the clinic from the "*faculty lounge"* that inhibited communication with colleagues. This theme will be discussed further.

The issue of time was a recurring theme and Chris further described how it impacted directly on assessment.

> *We don't have <u>time</u> at end of term to go over this. At <u>midterm</u> yes. We have an hour blocked out to allow us to go over. (Chris)*

While "*time*" had been allocated to the "*mid-term*" feedback, there was no time given for the supervisor to meet with the student to discuss the final assessment. This assessment was the one discussed earlier that was in a tick-box on-line format. A lack of time given to a task devalues that task, and hence encourages the tendency, I propose, to award an 'A'. In another institution time or the lack of it was also impacting on the value given to the instrument.

> *If we want to make it very formal, serious about assessment we need <u>time</u> we need to have maybe one student one hour they do the consultation with patient we go with them from the beginning to the end, not we are busy with the other students and the patients we cannot go from the beginning. We let them do the consultation first and in general what we think of them to give the mark not really at that moment. We need to have enough <u>time</u> to give a proper assessment. (Julie)*

To follow the formal procedures of the assessment process needed dedicated time to be allocated. Not in this case the problem of sessional work, it was the daily clinical structure within which the supervisors operated that prevented this. Govaerts *et al.* (2013) also discuss time and situational constraints as factors impacting on assessment in the workplace and propose that it actually contributes to raters using pre-existing (personalized) schemas.

In the smaller institution time was allocated for meetings between supervisors and managers regularly. Further, supervisors had developed a group e-mail communication system to discuss any issues coming up in clinic.

> *We also can <u>email</u> as a group, to highlight we have a problem, and warn everybody, especially if this person [student] is booked on different days, seeing different supervisors. So it's sort of a red flag for all the supervisors to keep an eye open for problems. (Kate)*

For this group of supervisors communication was presented as a major support for their role.

6.7 Identity

The general background of the supervisors was varied with supervisors having

been in practice from 5-38 years, employed as supervisors for 2-22 years and 50% (7/14) of the supervisors had graduated from their own college. All worked part-time and maintained busy private practices. In line with the previous discussion on the lack of formal training of supervisors, only three of the fourteen supervisors interviewed had had any teacher training, none in clinical supervision specifically and none in the use of assessment instruments. Two supervisors described training in previous jobs that they felt had been relevant to their current role as supervisors especially in the skill of giving feedback to students. Also, as discussed previously, any in-house training to be a supervisor was mostly very brief involving a basic orientation to the clinic and shadowing another supervisor for usually as little as a day; only one supervisor described an extended period of shadowing an experienced supervisor. It must be noted that 10 of the 14 interviewed had been working as supervisors for over 8 years and so were highly experienced. In line with the findings of Bloxham and Boyd (2007) and Rust (2007), nine supervisors described modelling their supervision style on their own experience as students.

Lacking training in the role of being a supervisor or the language of competencies and lacking involvement in the development of assessment instruments led, I argue, to supervisors utilising the skills they had which were as practitioners of Chinese medicine trained in patient care. That their own personal understanding of supervision was gained primarily from their experience as students themselves supported this. The 'authentic' structure of the teaching clinics mirroring a CM practice, I propose, perpetuated the primary identity of the supervisors as practitioners of CM rather than teachers. A few supervisors relayed that they discussed issues of clinical supervision with CM colleagues outside of the institution because of lack of time and the busyness of the day when in clinic. This would add to, I argue, the supervisors' identifying as practitioners as opposed to teachers working within an educational institution. In line with the research of Boyd and Bloxham (2014) their subject discipline identity dominated.

Supervisors had personal perspectives on what they wanted to impart to students. One supervisor was keen to teach students how to be "*busy*" and "*run*

multiple rooms" and "*multiple patients*". The supervisor was very aware of the debt students would graduate with and wanted to ensure they could earn a living. Many expressed awareness that they were training students for their future careers as CM practitioners and so wanted to "*show how CM operates*". Mitch talked about allowing students to develop their own style.

> *and if I am constantly dictating how I use the instrument* [the practice of CM] *then they are not going to find that kind of sweet spot they need until after they graduate and it is going to take longer.*

This extract displayed a definite distancing of this supervisor from any standardized outcomes. While pedagogical methods of supporting students' learning were inherent in the interviews, as presented earlier, supervisors did not present as seeing their role as teachers enmeshed in any deeper pedagogical understanding of WBL or WBA. Their identity could be seen as that of mentor inducting students into the practice of CM. Remember, this was also the opinion of one of the managers as presented in the previous chapter. Thus the dominant identity of the supervisors was in the discipline of CM not the discipline of education.

The institution in the UK did illustrate a different approach to developing a joint communal identity as teachers developed through their calibration meeting.

> *and all of us have very different <u>styles</u> how to maintain the curriculum, we follow the same curriculum, so that's quite a constant effort I think. We are all quite successful long term experienced practitioners out there, and we are all <u>different characters</u> out there, how to work under one roof, under one curriculum, hold the same standard, and constant negotiation, recalibrating, learn from each other, argue with each other… (laughter). (Kate)*

With "*different styles*" of the practice of CM, as well as their "*different characters*", Kate described how this led to difficulty in maintaining a common curriculum with standardized learning outcomes. It was not the learning outcomes that created the shared mental model but that time to discuss, negotiate and argue moved the competencies being applied away from staying personalized and thus open to bias. Further, the supervisors had articulated differentiated criteria in their assessment instrument to describe a standard of

performance, and so had mutually developed the language in the assessment instrument. This resulted, they proposed, in the ability to give differentiated grades. In the other two institutions differentiated criteria were not evident and the supervisors had not been involved in developing the assessment instruments. While Steinert, O'Sullivan and Irby (2019) acknowledge that there has been a lack of research demonstrating the effectiveness of staff development activities on teacher identity formation, they do present learning on the job and also belonging to a community of teachers as approaches for developing a shared identity.

Summary

The discussion above has gone some way to answer my research question: 'How is CBME being enacted in the institutions (supervisors) and what issues does this present?' The supervisors were very aware of their role in educating students, in developing students as competent practitioners who would be effective, safe and care for their patients. Their focus, however, was not on assessment, if that meant the production of a discriminating grade. The supervisors were applying personalised competencies to a psychometric assessment instrument that produced a measurement, the mainstay of CBME, but which did not, I propose, have validity or reliability.

For the supervisors, assessment was ongoing and in the moment, assessment-for-learning rather than the application of the learning outcomes as presented on a psychometric assessment instrument. Assessment-of-learning, of 'warranty', took a secondary position. 'Failure-to-fail' and 'grade inflation' was linked to the intimate supervisor-student relationship in which grades were given to encourage students not differentiate their abilities. Ball (2017) considers that supervisors' judgements are often superceded by demands of measurement, but in this research this was not in evidence. For the supervisors, learning outcomes and assessment instruments did not dominate. Assessment had been bent and changed to allow the supervisors to relate to students and their environment in a way that made sense to them and the culture and style of

practice they identified with. This was not an explicit rejection of the learning outcomes but a 'working round', a 'principled infidelity' (Hoyle and Wallace, 2007) or what Gleeson and Shain (1999) call a 'strategic compliance'. The role that supervisors took was related to: the lack of objectivity possible in authentic settings, the intimacy of the work-place, 'knowing' students and a lack of training in the assessment instruments. It was communication between supervisors and with managers, or lack of it, that determined the possibility of inter-rater reliability and a shared mental model.

Conclusion of Chapters 4, 5 and 6.

In line with the views of Govaerts, van der Vleuten and Holmboe (2019), this analysis has provided evidence of the tensions within a CBME approach to assessment when standardization and psychometric assessment tools come up against the authenticity of the work-place and when multiple atomized competency statements are spun together with personal judgement and the need to assess WBL holistically. The entangled phenomenon described above of three stories of WBA and how they manifested the constructs of CBME, however, were not easily explained by the different approaches of the instrumental-technical and socio-constructivist and the different roles and responsibilities of the stakeholders. My analysis needed to move away from the linear nature of the end-directed instrumental action seen in WBA (Carr, 2005; Gherardi, Nicolini and Yanow, 2003) as presented in CBME to a view of assessment as a 'practice'. In the next chapter, 'practice' will form the framework for a different discussion about WBA.

Chapter Seven: Viewing Workplace-Based Assessment as a Practice

My analysis so far has led me to consider that the assessments being carried out were being enabled and constrained by discursive, material and social factors that the tensions within CBME, as a theory, could not explain. By the application of practice theory to my analysis I was better able to make visible the inter-relationships and entanglement of the stakeholders and the artefacts and spatio-temporal activities and arrangements that formed the practice of WBA. Mahon *et al.* (2017) pose useful questions when researching practices and arrangements that helped frame my analysis and identification of empirical connections. What were the consequences of the nexus of sayings, doings and relatings of all the stakeholders; what local arrangements were shaping assessment as a practice in the institutions; what traditions were in operation for example CM or CBME? These questions provided a better purchase on understanding what was enabling or constraining the practices of assessment.

7.1 Practice as sayings - cultural and discursive

As discussed earlier, a practice is constituted by what people think and say in words spoken and written. The language chosen in the assessment instruments was communicating specific meanings about CM knowledge and practice and how it was to be assessed. The competencies were written in a standardized and atomized format that leaned towards the measurement and accountability requirements of accreditors in response to the dominance of CBME. Transposed into learning outcomes, some had been re-aligned to the competencies of a different health profession with sometimes little adaptation to a CM discourse. The discourse that spoke to the philosophy and principles of CM, while present in the general Mission Statements and Clinic Handbooks, had been stripped out of the learning outcomes. In the translation and codification of competence to competency statements made by the professional bodies to learning outcomes, important representations of the practice of CM had been left out. Meanwhile, the supervisors, not focused on the published learning outcomes were applying their personalized competencies to the WBA and exercising their own discretion as to what skills and attributes held value.

This did not seem to be the result of a misunderstanding of the outcomes or the 'creative insubordination' described by Schubert (2008, p. 410) but more in line with what Freidson (2001) considers intrinsic to how a professional goes about their work. The published discourse of CM held less meaning than the supervisors' own professional and personal experience, their embodied knowing of CM. This clash of discourses, that of insight, devotion and a loving and caring heart, that emerges from what Young (2013) calls a more 'sacred' tradition, and one of objectivity whose aim, he continues, was to free itself from the constraints of the 'sacred', impacted on the practice of WBA. The supervisors' different interpretations of CM had not been translated into the official learning outcomes, and yet it was these that produced the grade. Tavares *et al.* (2019) recommend that stakeholders need to recognize their assumptions in the assessment of clinical competence.

> Assessment is optimized when its underlying philosophical position produces congruent, aligned and coherent views on constructs, assessment strategies, justification and their interpretations (p.1).

I have shown how the very format in which the instruments were written with their linearity and atomization and numbers of outcomes, were also leading to a devaluing of the instrument. Supervisors reported how they applied an holistic approach to assessment and did not go through each atomized learning outcome to come to a grade. The psychometric discourse of 'warranty' of the instruments that assumed objectivity could not be upheld by the supervisors who were immersed in the subjectivity of the triadic relationship between themselves and the student and patient.

The discourse of education or of CBME was also one that most supervisors had no training in along with little induction into the meanings and interpretations of the assessment instruments. Their main understanding of assessment had come from their own experience as students. Thus again, their practice of WBA was rooted more in the discourse of CM. Blind to the epistemologies inherent in the assessment instruments, any situations arising in the assessment process were interpreted, understood and acted on based on their identity and

embodied knowing as practitioners of CM. Thus, I propose, the legitimacy and use of the instruments were being constrained by the clash of discourses.

There are issues, however, in applying personalized competencies to assessment due to the potential for these to remain implicit or unconscious on the part of the assessor (Orr and Bloxham, 2013), have not been made explicit to students (Brooks, 2012; Sadler, 2009) and have not been agreed amongst all the assessors (Jawitz, 2009; Meddings, 2017). For two institutions, the published learning outcomes were not a result of any 'shared mental model' arrived at through a discussion between the writers of the instruments and those enacting the instruments or amongst themselves. The outcomes the supervisors were applying came from a personalized discourse that remained tacit. In the third institution a shared discourse had come about through their mechanisms of communication - the email group, the development of assessment criteria and their calibration meeting at which, as Kate described, they *"learn[ed] from each other, argue[d] with each other"*. For Price *et al.* (2011) 'assessment standards reside in academic/professional communities' (p. 484) and 'how the assessment environment is managed impacts strongly on the effectiveness of assessment' (p.488).

More than 'sayings', I would support Gerhardi's (2009c) position that incorporates all the senses in developing an explanation of WBA as a practice as it is through the body that we mediate with the world. The concept of embodiment, I argue, resonates with my analysis of WBA. Supervisors talked about the intimate relationship that developed between themselves and the students, working in direct relationship with the patient. They described being in harmony with their students, '*knowing'* their students and thus they did not need a cognitive understanding of the learning outcomes within the assessment instrument to derive their judgement of the student's competence. Just as in the CM diagnostic tradition meaning about patients is gained through the senses - seeing, palpating, hearing, smelling - so I interpret that it was the supervisor's body with all its senses that had become the locus of the practice of WBA. Not confined to the mind alone practice includes the whole professional as an embodied being involved in a shared and meaningful world. What has been

revealed as important is that this embodied and tacit understanding of students'
competence was finding its way into the marks the supervisors attributed, as
quoted by one: "*I give them one or two marks anyway*".

7.2 Practice as doings - material and economic

A practice is constituted of what individuals do in a physical and material space-
time and within an economic context of resources, production and exchange.

While the value of the WBA might have been low in terms of a student's overall
degree grade with assessments that were more objective in their structure
being given more emphasis, the importance of the workplace was evident
through the resources incurred by the material and human facilities the
institutions provided. The students were immersed in the day-to-day work of the
clinic with supervisors and the administrators such as receptionists and
technicians in support. Supervisors worked closely with the students as they
treated patients, accompanying them in their treatments, checking how patients
were positioned, checking point locations through palpation and correcting
needle depth and angle. They made sure students were dressed appropriately;
hair tied back, fingernails clean. They guided students in the immediate care of
patients carrying out assessment-for-learning. The materiality and
paraphernalia of the clinic with its patient waiting rooms and treatment couches
and white coats and needles and acupuncture charts, were, I argue, critical for
the enactment of being a practitioner of CM for students and also supervisors
and thus for the practice of WBA. What I am proposing is that WBA was
embedded in the integrity of the clinic as a place where CM was enacted, in the
routines, protocols and artefacts (Evans and Guile, 2012a) that framed the
student's performance and in the role modelling and mentoring from
supervisors themselves members of the wider CM community. Kemmis and
Grootenboer (2008) describe how non-human objects and artefacts are not just
a backdrop of practice but a dimension of it with human actions and work
'hanging together' through material arrangements including human bodies. It
was the material activities and arrangements of the clinic that constituted WBA
as a practice. This is what Shulman (2005) describes in professional education
as 'signature pedagogies'. What Shulman means by signature is a way of

teaching (and assessing) that is 'routinized', 'distinctive in that profession' is 'pervasive within the curriculum' and can build 'identity and character, dispositions and values' (p.9). Signature pedagogies are habitual and routine, accountable through being visible, interdependent and collaborative, uncertain and unpredictable, and emotional and affect-laden. Shulman proposes the clinical rounds in hospitals in medical education as 'signature'. In this research the teaching clinics are the basis of the 'signature' pedagogy, leading to 'signature' assessments.

The physical location and structure of the clinics were also materialities that enabled or constrained the possible inter-relationships between supervisors, managers and other academics. The smaller institution displayed a more developed relationship between supervisors and management made possible, I propose, by the intimate nature of the building as described earlier with their common social space leading to regular informal communications. In the two institutions in which the clinic was more distanced (necessitated by the size of the institutions and access needed by the public) from the administration and teaching rooms and the managers developing the instruments, the supervisors were manifesting more problems with the assessment instruments either in their lack of valuing of them or in their actual use. In one institution, with limited times to meet, supervisors could carry out their supervision duties only coming into contact with the dedicated clinical administrators and only briefly with other supervisors who might be in on the same clinic shift. The busyness of the clinic was presented in another institution as a factor that inhibited the formal process of assessment. Thus, in my analysis of these three institutions, the architecture and physical location as materialities were impacting on the practice of WBA.

The assessment instruments were a further materiality. Managers and supervisors saw the input of hours along with the feedback supervisors gave in relation to each patient treatment, as the primary evidence of a student's competence. The assessment instruments seemed to be a very minor aspect of this embodied world. If the student "*turned up*", "*followed the rules*", "*cleaned the rooms*", "*did all the paperwork*", then they would get an 'A'. As discussed, the very value of the assessment instrument was being undermined by a

multitude of factors from not just the supervisors but also the accreditors and institutions. The linearity, numbers of learning outcomes, confusion of format and the stakeholders' different understandings and descriptions of competence all constrained the value of the assessment instrument.

Gherardi (2009b) presents a view of materiality that offers another explanation of these aspects of materiality, the clinic architecture and the assessment instruments.

> … practices are stabilized so as to restrict the space of possible negotiation and to inscribe certainty of performance in the practice itself. Practices form a role of uncertainty reduction when the preconditions for their performance are anchored in a material world made of objects, tools and technologies which direct actions and forbid undesired ones. (p. 356)

The relevance of this statement in this context was that the institutions, in their structure and processes played a normalizing role (Hager, 2012) in the maintenance of the practice of WBA. The space of negotiation was limited by the architecture of the clinic, the students' and supervisors' clinic timetables, the comings and goings of the patients, the routines to follow and forms to fill in. Certainty of performance could be seen in the way supervisors worked with students following the systems within the institution, systems that many had worked with as students themselves. While the architecture constrained it also enabled the practice of WBA.

At the same time as the format of the assessment instruments with their standardized learning outcomes were being devalued they were, I argue, a critical component of the practice of WBA. Without this artefact, the formal act of assessment could not happen for the institution and ultimately the student no matter how lacking in objectivity the marks placed on it might have been; a grade had to be given and was given. As stated, the strong normative dimension of practices could be seen in the enactment of the filling in of the instruments. An unexamined agreement was being maintained between the institution's need to record a grade and the supervisor's role to grade, even though the competencies they each aligned with were quite different. As presented earlier, the managers saw the value of the formal assessment as

more to encourage than measure and the acceptance of grade inflation by managers and supervisors was a tacit rule. Thus, even though the assessment instrument constrained the practice of WBA, the intentional directedness of the action of filling it in is what enabled the 'doings' of WBA to hang together.

Thus the practice of WBA in these institutions was constituted by its 'doings', with the materiality of the clinics both enabling and constraining WBA.

7.3 Practice as relatings - social and political

A practice is constituted by its 'relatings' historical, current and anticipated (Kemmis *et al.*, 2012) formed between people - the teacher and student and including the patient, the managers and the teachers, the institution and the accreditor - and between people and artefacts.

People's roles and routines and positions of legitimacy and power determined their values and the values of the artefacts they were connected with. It was the managers who had created the assessment instruments and thus they held value within the institution. As relayed by two supervisors, the position that the managers were taking in aligning the WBA with CBME, was seen as valid in relation to the external stakeholders, the accreditors and the desire for status of the institutions (and I add economic pressure) to be in a wider HE and medical world. Looking to the future, the institutions saw this alignment as a way to legitimize CM. The accreditors, institutions and the supervisors were working in a world where accountability and standardization have now become a powerful force. While it was the accreditors and the institutions who were experiencing this 'governmental' force, the supervisors, while working locally, were tacitly supporting this situation.

The supervisor's dominant relating to WBA came through their identity as practitioners of CM and participation in the carrying out of CM; remember only three supervisors had any teaching qualifications and it was not an expectation of two of the accreditors or institutions. The spatio-temporal structures in which WBA was embedded was the clinic with its artefacts and systems and rules, in

which patients came and went. Very little time was given to the formal undertaking of assessment and reflecting on the learning outcomes. It was their relationship with the materiality of the clinic that created the sense of agency and purposiveness for the supervisors and determined their practice of WBA. As described, the supervisors could be seen to alter the assessment format to better mirror the way they worked with students. Lacking training in education and induction into the atomized knowledge contained in the learning outcomes the tendency was for the supervisors to apply their subjective, interpretive, inconsistent and intuitive understanding of professional work. Further, it was through the acting out of their personalized competencies onto the assessment instrument, and being allowed to implicitly by the managers (for example, in two institutions there were no differentiating criteria to explain the values of the individual outcomes and the 'A' award was condoned) that agency was maintained by the supervisors.

The practice of assessment in all the institutions as shown through 'grade inflation' and 'failure to fail' could, I propose, be the direct result of the intimacy of the relationships between supervisor and student. The supervisors were involved in assessment-for-learning not warranty, not 'mining' for whether the student met each of the learning outcomes but 'travelling' with the student as they became part of the community of CM. As shown in the analysis the supervisors' awareness of the students' personal lives and the debt they were incurring showed that these infiltrated their dispositions and actions with the marks allocated in the WBA seen as a validation of the relationship.

Moving beyond the local, employment law in each country also determined the relationships possible between the institutions and the supervisors. In the two institutions in which the supervisors had tenure (a legal employment requirement), there was more time available for meetings and a sharing of ideas. In one of those, however, the busyness of the clinic intruded on that availability. In the institution with supervisors on smaller numbers of hours (remembering that this was a desirable condition for the institution and the accreditor), there was awareness that this inhibited communication and the development of a shared relationship with WBA. Differences in employment, I

propose therefore, were factors impacting on the availability of supervisors to communicate with each other and participate in general curricular developments and thus the development and implementation of the institutions version of WBA. 'Time' and or 'lack of time' brought about by the employment structure, was a pre-condition for the successful development of relationships and subsequent enactment of WBA.

The psychometric marks annotated on the assessment instruments were the manifestation of entwinement and inter-relationships, not just the product of the supervisors' own agency but the exigencies of the situation (Kemmis and Grootenboer, 2008) and the contradictory and contesting ideas and relationships of the different stakeholders.

7.4 Five types of Relationship

In trying to show the intimate relationships between practices and material entities, Schatzki (2012; 2016) has described five types namely: causality, constitution, intentionality, pre-figuration and intelligibility. Considering these more closely has helped me to gain more analytic purchase on how practice theory can help explain WBA.

'Causality' refers to activities that affect and in turn respond to arrangements not so much as a linear relationship but in their entanglement (Fenwick, 2014). This could be seen in the location of the clinics that distanced the supervisors from the managers, and the employment contracts, having a causal relationship in enabling or constraining communication. Further the busyness of the clinics and lack of time allocated to devote to the assessment instrument, constrained the ability of supervisors to come to a truly considered grade for students.

Another example of causality, I propose, but more tentatively, was illuminated by how supervisors described their appreciation of their students. The extracts cited in the previous chapter – that the supervisors interpreted the student's choice of them as being "*lucky*", that students who chose them were "*all stars*", that this choice allowed for a "*harmonization*" of energy and a positive

relationship – manifested a perception by the supervisors that they had been personally selected by the students. This, I argued earlier, was one of the rationales for awarding high marks. This situation could, however, be interpreted as a combination of the booking structure of two of the institutions that allowed students to "*choose*" their supervisor along with the possible discrimination against a supervisor that could lead to their not being employed. If a supervisor awarded low marks subsequent students might not choose them as supervisor. Thus, the system of the clinic bookings along with the type of employment contract, had a causal relationship to the practice of WBA as shown in grade inflation and failure-to-fail.

By 'constitution' is meant activities and arrangements being essential for each other. Without the physical structure of the clinics with their artefacts, both objects (waiting rooms, beds, needles) and humans (patients, students, supervisors, receptionists, managers), the practice of WBA could not happen for the supervisors or the students. A further example was the assessment instrument that was constitutional to the allocation of marks however they were reached. The accreditation process itself was also constitutional, essential for the existence of the institutions but in different ways. In Australia and the USA there was a legal requirement while in the UK it was the desire to improve the status of CM that directed the institutions in choosing CBME as their curricular model.

'Intentionality' concerns the directedness of activities and thoughts and imaginings of participants towards arrangements. Dunne (2005, p. 368) sums up a definition of practice presented by Alasdair MacIntyre as:

> a coherent, complex set of activities that has evolved cooperatively and cumulatively over time, that is alive in the community who are its practitioners, and that remains alive only so long as they remain committed to sustaining – and creatively developing and extending – its internal goods and its proper standards of excellence (this commitment constituting them *as* a community)

My interpretation of this statement was that it was the intentionality or the 'purposiveness' of the practice of WBA to develop practitioners of CM, which aligned the supervisors, accreditors and institutions. Intentionality could be seen

in their commitment to the promotion of CM however differently that was understood. Further, the employment of working practitioners as part-time supervisors sustained the practice of WBA, albeit a different WBA practice to that of the accreditors and managers.

In his quote above, MacIntyre proposes that to remain alive a practice needs to not just sustain but creatively develop and extend its internal goods. This statement is relevant when I further tried to understand why there seemed to be an obstruction with the two institutions who were trying to implement a new assessment instrument. Two factors appeared as important. Firstly, I propose, the lack of understanding of CBME was due to the marginalising of the discipline of education combined with the lack of induction into the assessment instrument. This latter was added to by the lack of time given for discussion of assessment issues due to employment conditions and the busyness of the clinic as described earlier. Secondly, the supervisors emulated the supervision and assessment process they had experienced as students with many also being graduates of their own institutions. As one manager reported, the supervisors were all 'on the same page' thus allowing for ways of working to be passed on with no further training. While this supported a continuity and a degree of cohesion amongst supervisors it could be seen to be leading to what Kemmis and Mahon (2017) call 'sedimentation'. Instead of remaining fluid the structures and practices and relationships had become institutionalised, they were no longer contested and so functioned as 'mediating preconditions' for further practice limiting change. While the personalised competencies of the supervisors and understandings of the assessment instruments remained tacit the introduction of new WBA processes or formats would be constrained.

As described earlier, intentionality was also manifest in the assessment instruments and the action of the supervisor in annotating it with a grade. The instrument could be seen to form a kind of infrastructure through its directing of the assessment activity and links us back to Gherardi's (2009b) proposal that practices inscribe a certainty of performance. At the same time, the objectivity of the instrument was constrained by the implicit rule of 'grade inflation' and 'failure-to-fail', a tacit intentionality.

By 'prefiguration' Schatzki is describing how arrangements have bearing on future courses of action. Practices give meaning to and affect arrangements and arrangements facilitate and prefigure practices. As the data has shown, the lack of implementation of the objective measurement of atomized learning outcomes was prefigured by a number of factors. Firstly, the implicit belief in the input view of the curriculum by accreditors, managers and supervisors was still buried beneath the measurement manifestation of CBME and thus could be seen to prefigure failure-to-fail. Students who completed their hours were deemed competent. Prefiguration could also be seen in the lack of training in education and short term or tenured contracts that did not (or did) allocate time for meetings and communication. As a result supervisors' identified as practitioners of CM inducting students into CM and not as assessors annotating the instrument for 'warranty'. The format of the assessment instruments and the introduction of a tick-box on-line process in one of the institutions also pre-figured the awarding of an 'A'.

'Intelligibility' refers to the meaning arrangements hold for participants. The assessment instrument had different meanings for the different stakeholders dependent on the regulatory demands and the aims and purposes of their roles. For example, the discourse of standardization and atomization enacted by the accreditation and institutional documents met their quality assurance role and alignment with HE and other health professions. The managers also had a quality assurance role to certify achievement for accreditors and students and at the same time to maintain a functioning workplace. For the supervisors promoting learning was their aim as they worked intimately with the student in patient care. It was by immersing the student in their world of CM that they assured them to go out into the world. The language of the competencies did not hold intelligibility for the supervisors in relation to the performance being observed. It was the different internal representations, their understandings of the world of CM and of assessment including their beliefs, habits, skills and interests and their own unique sets of rules that determined how stakeholders understood and enacted WBA.

7.5 Webs of Possibilities

The multiple factors described above could be seen to create what Kemmis and Grootenboer (2008) call 'webs of possibilities' or Fenwick, Nerland and Jensen (2012) 'systemic webs', that enabled or constrained the practice of WBA.

As I have shown, WBA was being enacted differently by the multiple stakeholders - by accreditors, institutions, managers and supervisors (also by students and administrators but those are beyond this thesis). Not just webs of possibilities within the institutions, Kemmis and Grootenboer (2008) describe how any one practice is the product of, and inter-related to, other practices, meta-practices that regulate, constrain, enable or influence other practices. Schatzki (2011) introduces the concept of the thickness or thinness of the relationships between practices to explain these webs. The accreditors had a thicker relationship with the practice of HE and CBME within their respective countries but only a thin relationship with the practice of CM as manifest in the content and format of assessment instruments that lacked its heterogeneity and multiplicity. Institutions in their support of authentic clinics and the employment of part-time supervisors who remained steeped in practice manifested a thick relationship with CM. At the same time their documentation showed a thick relationship with the practice of HE and the accountability role of accreditation. Supervisors' relationships were primarily with the practice of CM and little with the practice of CBME. Their relationship with the assessment instruments was fleeting, only forming when a grade was needed; measurement was not a supervisor's concern. The assessment instruments were aligned with the practice of CBME while the assessment-for-learning was aligned with the practice of CM.

The following is a developed example of the complexity that these webs can form and explains how WBA was being enabled and constrained differently in different institutions by discursive, material and social factors. As discussed before, the smaller institution manifested a conscious, consistent and more objective application of the assessment instrument. This emerged from the purposeful calibration meetings of the supervisors and managers that allowed

for the skills of supervision and of assessment to be learned through marking and moderating as a team. The following factors seemed to create this possibility. Firstly the practice of employment law in the UK required that if a person had regular employment, even if part-time, they must have tenure. This led to fewer supervisors being employed on more long-term contracts, embedded meeting time and thus promoted on-going regular communication. Secondly, the practice of accreditation and validation as a result of regulatory structures in the UK required supervisors to give a rationale for their judgements to an external examiner. That external examiners scrutinised assessments and challenged a supervisor's marks, was proposed by one manager as a factor in helping develop the role of supervisors as assessors. This too drew the supervisors and managers together. Further, the shared professional learning gained from the involvement of supervisors as external examiners or on accreditation boards for other institutions was also cited as allowing for the assimilation of the discourse of CBME. The impact of these meta-practices came together to stretch the identity of the supervisors more towards that of educators allowing for the enactment of a more objective WBA. The other two institutions did not have any regulations requiring external examiners and as a result there appeared to be little external scrutiny of assessment processes and so little opportunity for supervisors to work with the rules both implicit and explicit of WBA.

Summary

WBA in the three institutions I have researched 'hangs together', I propose, through the common purpose of the different stakeholders to induct students into the world of Chinese medicine. How that was to be done and how the construct of CBME was managed, however, differed depending on the sayings, doings and relatings that comprised each stakeholders' particular assessment practice. As we have seen the grades being entered by supervisors were not the result of the linear application of observed learning outcomes to the assessment instrument but of multiple factors and the entwinement of all the stakeholders. The accreditors' enactment of WBA was reaching out to the wider HE and healthcare community focusing on measurability and accountability,

promoting and upholding the status of CM but using a language that had lost the discourse of CM. The institutions were leaning out towards the accreditors' mapping competencies onto learning outcomes and providing authentic teaching clinics in which to enact WBA but at the same time leaning into the reality of that authenticity, condoning grade inflation and failure to fail. The supervisors, with their personalized competencies, identity as practitioners of CM and intimate relationships with students and patients focused on assessment-for-learning but supported the formal WBA system through the act of filling in the assessment instrument. For Reckwitz (2016, p. 394) 'every social practice has a built-in affective dimension' and it is through motivation (what Schatzki includes under the teleo-affective) that affect comes into play. Participation in a practice always involves some affective incentive to participate, as noted by the manager who proposed that the supervisors *"want to be connected to a community and give back, they are not disgruntled"* (Mavis). Reckwitz continues that it is architecture and artefacts that can act as 'affect generators' within practices' (p383) creating the 'mood' of the practice. Viewing WBA as a practice, with its emphasis on the relationships between the socio-cultural-material has shown the inseparability of the assessment process and product and those enacting the assessment, the accreditors, institutions, managers and assessors, artefacts and material arrangements. Any changes to the practice of WBA would need significant attention paid to these local enactments.

Chapter Eight: Conclusion

In this chapter I return to my original questions and summarise the evidence to support my conclusions on how WBA was being enacted by the different stakeholders. I will highlight the theoretical contribution of this study to WBA and review both the methodological innovations that made this possible and outline some of the limitations of the research by recommending further research areas. The empirical evidence gained from these three cases is a first attempt to map out a small part of the terrain of WBA in Chinese medicine. As explained in my methodology, given the limited number of institutions within the study, any generalisation from this research is best considered as a working hypothesis rather than a conclusion. Thus it is for the readers of this work to consider if the evidence and my explanations resonate with their own experience and in that I believe my results have implications for the wider health-care professions.

The accreditors' primary manifestation of WBA was of an instrumental-technical discourse as shown in their drive for standardization and measurability in line with HE requirements. While they called for authentic places for learning and assessment this was not tempered with how that contextual, situationally determined and tacit knowledge of CM might act on WBA, the training of supervisors, or communication needed for this realisation. I argue that, in their striving for external legitimacy, the discourse of CM had been stripped out. To gain status in a political situation dominated by orthodox medicine, scientific referents have replaced those of CM (Cant and Sharma, 1996; Glatz, 2019) and there has been a tempering of holistic knowledge claims (Givati, 2015). The discourse of standardization and atomization was carried forward by the institutions. Focused on mapping and measuring, there was little acknowledgement of the discourse of CM in their published learning outcomes. At the same time, it was the major resource of the architecture and artefacts of the authentic environment, along with the tacit rule of 'grade inflation' and 'failure-to-fail' that promoted learning over assessment and the discourse of CM. For the supervisors, steeped in the practice of CM, with no training in the

constructs and values of CBME and immersed in the triadic relationship with the student and patient, WBA was ongoing and in the moment. It was the discourse of CM that dominated their life-world, the measurement discourse only manifest in the filling in of the assessment instrument.

It was the misalignment between how assessment was described by accreditors, interpreted by institutions and carried out by supervisors that illuminated the enactment of WBA, made possible by the case study method employed. And it was practice theory, with its emphasis on the socio-cultural-material activities and arrangements that sensitized me as a researcher and provided a framework to understand how different assemblances lead to those different enactments of WBA. CBME as a construct, I conclude, is under-theorized even when both the instrumental-technical and socio-constructive approaches are taken into account.

The result of my using practice theory as the lens with which to view my research has allowed me to recognize a fundamental flaw in the construct of CBME. Drawing on Dall'Alba and Sandberg (2014) I argue that the very need to strive for impartiality and distancing following a CBME construct assumes not just a recognition of context as in the socio-constructivist approach, but of entwinement. Developing this argument, a further assumption embedded in CBME is explained by Carr's (2005) definition of a 'practice' (in this research the practice of assessment).

> a practice cannot be understood as an instrumental means to some predetermined 'end', because the 'end' of a practice is internal and inseparable from the practice that constitutes the means of its achievement' (p.337)

My interpretation of Carr is that any understanding of WBA needs to move beyond seeing the focus on the tasks the student performs as the end. The results of my research allows me to propose that WBA in programmes of study that lead to professional entry in Chinese Medicine (and I add in any professional entry programme) takes place within a specific practice world constituted of the entwinement of the cultural-discursive, material-economic and social-political arrangements, circumstances and conditions of all the

stakeholders, the accreditors, institutions, managers and supervisors. It is the different discourses, spatio-temporal arrangements, activities, artefacts, teleoeffective structures and understandings that are the enabling and constraining factors on WBA.

I would like to expand on two aspects from the research that I consider particularly significant for the development of the CM community. Firstly is the role of discourse especially how the CM discourse seems to be being sidelined by the dominant discourse of CBME in published documents. Secondly is the role of communication and how to support the development of an educational discourse within the community of practitioner-assessors. My focus on these aspects comes from my professional experience as an accreditor, manager and supervisor as described in the introduction to this thesis, that has given me an awareness of the multiple perspectives that the profession needs to work with. Ajjawi (2017) and Boud (2017) point out that the drive for public accountability is here to stay and thus the adoption of competency frameworks and statements is inescapable. Whitty (2008) accepts the position of other stakeholders in curricula development and proposes a collaborative and democratic approach to change normative ways of working. Leading educationalists on the BAAB also suggest that no profession can stay still and must constantly modify its norms to absorb current practices and accepted orthodoxies or consciously reject them (Hopper and Parrott, 2005). At the same time is the concern that the increased standardization and codification of CM is constraining the dialectic between the canons and the practice, the sacred tradition and artistic interpretive autonomy. Further, what is being enacted by the supervisors who carry out the WBA must be acknowledged.

8.1. The power of discourse

Chinese medicine has, over the last two centuries, in its drive for scientization and modernization, been transformed. Not just in the PRC but also in the West, the profession has absorbed new educational standards and accountability procedures. Professional bodies have translated the desired competence of its members into competency statements that have been interpreted by the

institutions into multiple learning outcomes to be enacted by the supervisors. Outwardly as written text, a standardized and professionalized concept of competence was in evidence in the three institutions. These disarticulated competency statements can be seen to be what Reckwitz (2016) calls 'inter-discourses' that is discourses that carry descriptive-normative representations of subjects, acting as the reference point for the curriculum and assessment and enabling and constraining actions and interactions. However, through gaining the voice of the supervisors, what this research has shown was that the written discourse of the competency statements that lacked a sensitivity to the practice of CM, might not have the power or authority assumed and did not constrain the supervisors' own expertise and interpretations of the practice of CM. For the supervisors, the discourse of CM was still holding onto its 'sacred' tradition and the different values that Ho *et al.* in their various papers have expressed. The dialectic between the canons and the practice was operating through the supervisors. The institutions could be seen to be practising what Givati (2015) calls 'pragmatic' holism, leaning out to a standardized and scientized interpretation of practice, but maintaining the professional integrity of CM.

The normativity assumed by the accreditors and institutions as being disseminated through the written texts, the discursive formation of a CBME version of CM as held within learning outcomes, are marginalized by the non-discursive, what Schatzki (2016) describes as the 'material-activity' of the supervisors. As has been discussed, standardized learning outcomes are not stable representations of knowledge either canonical or practical and they do not carry agreed-upon meanings, but come out of and are enacted by people and artefacts operating within complex environments and relationships. As the research of Tavares *et al.* (2020) concluded, there are competing social and practical factors that displace validity particularly arising from assessment judgements that need to be made in clinical practice. Thus can be seen how WBA is simultaneously manifesting both standardization and tradition. The challenge for accreditors and institutions is how to write 'meaningful statements of learning outcomes at a suitable level of abstraction' (Boud, 2020, p. 8),

learning outcomes tempered by an appropriate language that 'honour the traditions of Chinese medicine' (Anderson *et al.*, 2014, p. 705), that lean towards the personalized competencies as enacted by the supervisors allowing for indeterminancy and artistic interpretation - and retain the status of the profession and the need for accountability.

8.2 Communication and Community

As discussed, teacher training and induction for supervisors was lacking, thus limiting an appreciation of the CBME discourse manifest in the assessment instruments. Kemmis and Grootenboer (2008) highlight how practice theory could help explain the problems this raises.

> What makes a complex practice like education or medicine distinctive is the content of sayings, doings and relatings characteristic of the practice, and the way sayings, doings and relatings are bundled together in the conduct of the different professions (p. 51)

This distinctiveness of practice implies, they contend, that the movement of sayings, doings and relatings from one practice to another will be fraught with conflict as any dispositions and actions are shaped by both the content and activities and arrangements of each distinctive practice and interrelationships within that practice. That movement could be between the practice of education and the practice of medicine (ibid.) or, I add, the practice of Western medicine and that of Chinese medicine, or practice as defined by accreditors or institutions and supervisors, or the simultaneous occurrence of these practices. As an example, Smith and Levett-Jones (2013) describe the distinctiveness of the education and health professions as not just due to different philosophies and cultures but also purposes, the former more focused on students and the latter on patients. As shown, the assessment practice of the accreditors, which was focused on standardized competencies, had difficulty translating to the assessment practice of the supervisors who had to mediate between students and patients and the environment in which they were working. The tension between practitioner and academic roles has been researched in other professions such as physiotherapy (Hurst, 2010) and nursing (Duffy, 2013). As in those examples, the practitioner-teacher nexus in this study was leaning

heavily to practitioner. It was the supervisors who held the expertise and authority of competence in CM with their knowledge embedded in and supported by the architecture, routines and expectations of the workplace. Both explicit and tacit understandings and ways of working were integrated and embodied by the supervisors whose expectations of students were manifest in their personalized competencies. My argument is that, if supervisors are not inducted into the discourse of education and of the CBME discourse of assessment instruments, and if there is no opportunity to communicate and meet as a community of assessors to develop shared mental models, the supervisors' expectations of assessment remain that of their primary discipline. The relevance and value of the assessment instruments are thus disabled. Watling, Ajjawi and Bearman (2020) propose that how an assessment instrument is used in practice, depends on the community using the tool and their view of the world more than the instrument itself.

In focusing on the enactors of assessment, Davey (2006) explains the importance of the socialization and acculturation process in acquiring and accumulating experience. It was the relationships that formed in the calibration meeting between assessors and managers in the smaller institution that gave access to the ideas behind the assessment instrument. Any objectivity in WBA, I argue, emerges from the community of practice rather than the instrumentality of the learning outcomes and it is this that legitimates the value of the instrument (O'Neal, 2016). For Price *et al.* (2011) 'assessment standards reside in academic/professional communities' (p. 484). It is through academic relationships that supervisors become 'sharpened' (van der Vleuten *et al.*, 2010) to the assessment instrument and the discourse of CBME, and accreditors, perhaps, 'softened' to a CM discourse. Those relationships also allow for an alignment between the institutions' and the accreditors' understanding of the professional competencies.

What must be recognised in any negotiations around WBA is that there is an inevitable 'jostling of cultures' (Kennedy *et al.*, 2015) as the different aims and purposes of the stakeholders is enacted. For Gherardi (2009c) the concept of

practices 'as ways of doing things together' (p.547) promotes communication and community. He considers that

> what makes possible the competent reproduction of a practice over and over again and its refinement while being practised (or its abandonment) is the constant negotiation of what is thought to be a correct or incorrect way of practising within the community of its practitioners (2009c, p. 536)

Through leveraging the lifeworld of the supervisors and their passionate attachment to CM, by applying a language not of an instrumental and technical logic but one in which the supervisors' (aesthetic) judgement, their knowing-through-the-senses, is given status, and by creating the spatio-temporal conditions for relationships to develop, a more (qualitative) objective assessment of students competence to practise CM might be realised. In this way we might be able to sail more smoothly between the rocky shoals and whirlpools as described by Eva and Hodges in the introduction to this thesis.

8.3 Limitations of the research and way forward.

In this final section I would like to draw attention to some of the limitations of this research and areas for future consideration.

Atkinson (2017) warns that one cannot rely solely on what people report they do but need to consider what people actually do in real-time encounters. My time in the institutions, while limited, sensitized me to the role of the spatio-temporal arrangements and the interactions among the staff. The next stage of this research would be to involve more ethnographic work and to observe and document what Atkinson describes as the 'ceremonial order' of the social interaction, the activities both spoken and acted of the assessors and managers, along with the artefacts and architectures that make their actions possible. Paraphrasing Goldszmidt (2017) a further research question would be: in what ways does the clinical setting assemble to shape WBA? As this would mean participating in both patient and student encounters any ethical issues would need to be carefully considered.

A voice lacking in this research was that of the persons conducting accreditations. My analysis was of accreditation documents only and what I

have not accessed was the life-world of the accreditors as they applied these standards to institutional documents and observations. A brief analysis of research into accreditation processes shows a tension between adopting an evidence-based approach and diversity of practice as discussed in Innes *et al.* (2019) in relation to chiropractic. A future line of inquiry would take practices and arrangements as the focus and question the dialectic between contextual diversity and accountability (Bates *et al.*, 2019). A recognition of how the purposiveness of the assessment instrument as artefact allows WBA to hang together might be a way forward but this artefact needs to contain a familiar 'jargon' (Schuwirth and van der Vleuten, 2019). A shared language of competencies can only emerge from the local level and the politics and philosophical leaning of the institution and the community in which they are embedded must be illuminated. Foregrounding the nature and power of the signature pedagogy of the institutions, the workplace, could, I propose, also give meaning to the process.

Due to the lack of research into CM education in the West, my research process has been to seek out empirical evidence and theoretical considerations from the realm of *xiyi* (Western medicine) and adapt it to *zhongyi* (Chinese medicine) education. Just as I have needed to move and apply the sayings, doings and relatings from one practice to another, I would suggest that the world of CBME might consider the literature on practice theory as a way to illuminate some of the problems being encountered that I discussed in the literature review. This research has helped illuminate how educational aims might be being enabled or constrained by different assemblances and enactments and thus a way forward for workplace-based assessment could be to leverage the socio-cultural-material.

References

ACAOM. (2016). *Accreditation Manual*. [Online]. [Last accessed 10/06/2016].

ACAOM. (2017). *Standards and Criteria Manual: Master's Level and Postgraduate Doctoral (DAOM) Oct 21st*. [Online]. Available at: http://acaom.org/resources/standards-and-criteria-masters-level-and-postgraduate-doctoral-daom/. [Last accessed 20/01/2018].

ACAOM. (2019). *Comprehensive Standards and Criteria*. [Online]. Available at: http://acaom.org/resources/comprehensive-standards-and-criteria/. [Last accessed 03/09/2019].

ACGME. (2017). *Accreditation Council for Graduate Medical Education common program requirements*. [Online]. Available at: https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRs_2017-07-01.pdf. [Last accessed 12/02/2019].

Ajjawi, R. (2017). 'Problematising standards in an era of increasing transparency'. *Medical Education,* 51 (9)*,* 885-887.

Albanese, M. A., Mejicano, G., Mullan, P., Kokotailo, P. and Gruppen, L. (2008). 'Defining characteristics of educational competencies'. *Medical Education,* 42 (3)*,* 248-255.

Allan, J., O'Meara, P., Pope, R., Higgs, J. and Kent, J. (2011). 'The role of context in establishing university clinics'. *Health & Social Care in the Community,* 19 (2)*,* 217-224.

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*,* 95-103.

Anderson, B. J., Kligler, B., Taylor, B., Cohen, H. W. and Marantz, P. R. (2014). 'Faculty survey to assess research literacy and evidence-informed practice interest and support at Pacific College of Oriental Medicine'. *J Altern Complement Med,* 20 (9)*,* 705-12.

Andrews, B. (2015). *The making of modern Chinese medicine, 1850-1960*: Honolulu : University of Hawai'i Press.

Atkinson, P. (2017). *Thinking ethnographically*. (First Ed.). Thousand Oaks, CA SAGE Publications.

BAAB. *British Acupuncture Accreditation Board*. [Online]. Available at: https://baab.co.uk/about-us/about-us.html. [Last accessed 03/02/2018].

BAAB. (2016). *Accreditation Handbook*. [Online]. Available at: https://baab.co.uk/downloads/key-accreditation-documents.html. [Last accessed 03/10/2019].

BAcC. (2011). *The Standards of Education and Training for Acupuncture*. [Online]. Available at: https://baab.co.uk/downloads/key-accreditation-documents/item/17-bacc-standards-of-education-and-training-for-acupuncture-seta.html. [Last accessed 01/02/2018].

Bagnall, R. and Hodge, S. (2017). 'Using an Epistemological Perspective to Understand Competence-Based Vocational and Professional Education'. In M. Mulder (Ed.), *Competence-based Vocational and Professional Education. Bridging the World of Work and Education* (pp. 125-144). Edinburgh: Springer.

Ball, S. J. (2003). 'The teacher's soul and the terrors of performativity'. *Journal of Education Policy,* 18 (2)*,* 215-228.

Ball, S. J. (2017). *The education debate*. (3rd Ed.). Bristol: Policy Press.

Barnes, L. L. (2003). 'The acupuncture wars: the professionalizing of American acupuncture--a view from Massachusetts'. *Med Anthropol,* 22 (3)*,* 261-301.

Barnett, R. (1994). *The Limits of Competence. Knowledge, Higher Education and Society*. Buckingham: SRHE and Open University Press.

Bates, J. and Ellaway, R. H. (2016). 'Mapping the dark matter of context: a conceptual scoping review'. *Medical Education,* 50 (8)*,* 807-816.

Bates, J., Schrewe, B., Ellaway, R. H., Teunissen, P. W. and Watling, C. (2019). 'Embracing standardisation and contextualisation in medical education'. *Medical Education,* 53 (1)*,* 15-24.

Baxter, P. and Jack, S. (2008). 'Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers'. *The Qualitative Report,* 13 (4)*,* 544-559.

BERA. (2018). *Ethical Guidelines for Educational Research*. [Online]. Available at: https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018-online. [Last accessed 08/07/2020].

Berendonk, C., Stalmeijer, R.E. and Schuwirth, L.W. (2013). 'Expertise in performance assessment: assessors' perspectives'. *Advances in Health Sciences Education,* 18 (4)*,* 559-571.

Billett, S. (2010). 'Learning through Practice'. In S. Billett (Ed.), *Learning through Practice. Models, Traditions, Orientations and Approaches* (pp. 1-20). Dordrecht: Springer.

Billett, S. (2012). 'Practice-based learning and professional education'. In J. Higgs, R. Barnett, S. Billett, M. Hutchins and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 101-112). Roterdam: Sense.

Bleakley, A. (2006). 'Broadening conceptions of learning in medical education: the message from teamworking'. *Medical Education,* 40 (2)*,* 150-7.

Bleakley, A. (2010). 'Blunting Occam's razor: aligning medical education with studies of complexity'. *J Eval Clin Pract,* 16 (4)*,* 849-55.

Bleakley, A. (2015). *Medical humanities and medical education: how the medical humanities can shape better doctors*: Routledge.

Bleakley, A. (2019). 'Invoking the Medical Humanities to Develop a #MedicineWeCanTrust'. *Academic Medicine* 94, 1422-1424.

Bleakley, A., Bligh, J. and Browne, J. (2011). *Medical Education for the Future. Identity, Power and Location*. Netherlands, Dordrecht Springer.

Bleakley, A., Brice, J. and Bligh, J. (2008). 'Thinking the post-colonial in medical education'. *Medical Education,* 42 (3)*,* 266-70.

Blinman, T. (2017). *Competency Based Medical Education is Wrong for Surgery*. [Online]. Available at: https://doi.org/10.31219/osf.io/kzrqp. [Last accessed 08/07/2020].

Bloxham, S. and Boyd, P. (2007). *Developing effective assessment in higher education: A practical guide.* Maidenhead: Open University Press.

Bloxham, S., Boyd, P. and Orr, S. (2011). 'Mark my words: the role of assessment criteria in UK higher education grading practices'. *Studies in Higher Education,* 36 (6)*,* 655-670.

Bloxham, S., den-Outer, B., Hudson, J. and Price, M. (2015). 'Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria'. *Assessment & Evaluation in Higher Education,* 41 (3)*,* 466-481.

Boud, D. (2000). 'Sustainable Assessment: Rethinking assessment for the learning society'. *Studies in Continuing Education,* 22 (2)*,* 151-167.

Boud, D. (2007). 'Reframing assessment as if learning were important'. In D. Boud and N. Falchikov (Eds), *Rethinking Assessment in higher Education. Learning for the Longer Term* (pp. 14-25). Oxon: Routledge.

Boud, D. (2009). 'How Can Practice Reshape Assessment?'. In G. Joughin (Ed.), *Assessment, Learning and Judgement in Higher Education* (pp. 29-43). B.V.: Springer.

Boud, D. (2017). 'Standards-based assessment for an era of increasing transparency'. In C. D, B. SM, C. CKY and G. R (Eds), *Scaling up Assessment for Learning in Higher Education* (pp. 19-31). Singapore: Springer.

Boud, D. (2020). 'Challenges in reforming higher education assessment: a perspective from afar'. [Online]. *RELIEVE*, 26 (1). Available at: http://doi.org/10.7203/relieve.26.1.17088. [Last accessed 04/09/2020].

Boud, D., Dawson, P., Bearman, M., Bennett, S., Joughin, G. and Molloy, E. (2018). 'Reframing assessment research: through a practice perspective'. *Studies in Higher Education,* 43 (7)*,* 1107-1118.

Boud, D. and Falchikov, N. (2007). *Rethinking Assessment in higher Education. Learning for the Longer Term*. Oxon: Routledge.

Boud, D. and Molloy, E. (2013). 'Rethinking models of feedback for learning: the challenge of design'. *Assessment & Evaluation in Higher Education,* 38 (6)*,* 698-712.

Boudreau, J. D., Cassell, E. J. and Fuks, A. (2018). *Physicianship and the Rebirth of Medical Education*. New York: Oxford University Press.

Boyatzis, R. E. (1998). *Transforming Qualitative Information. Thematic Analysis and Code Development.* California: Sage Publications.

Boyd, P. and Bloxham, S. (2014). 'A situative metaphor for teacher learning: the case of university tutors learning to grade student coursework'. *British Educational Research Journal,* 40 (2)*,* 337-352.

Boyd, V. A., Whitehead, C. R., Thille, P., Ginsburg, S., Brydges, R. and Kuper, A. (2018). 'Competency-based medical education: the discourse of infallibility'. *Medical Education,* 52 (1)*,* 45-57.

Braun, V. and Clarke, V. (2006). 'Using thematic analysis in psychology'. *Qualitative Research in Psychology,* 3 (2)*,* 77-101.

Brinkmann, S. and Kvale, S. (2015). *Interviews. Learning the Craft of Qualitative Research Interviewing*. (3rd Ed.). Los Angeles: Sage.

Broad, B. (2000). 'Pulling your hair out: Crises of standardization in command writing assessment'. *Research in the Teaching of English,* 35 (2)*,* 213.

Brooks, V. (2012). 'Marking as judgment'. *Research Papers in Education,* 27 (1)*,* 63-80.

Broudy, H. S. (1972). 'Didactics, Heuristics, and Philetics'. *Educational Theory,* 22 (3)*,* 251-261.

Brown, N. (2000). 'What are the criteria that mentors use to make judgements on the clinical performance of student mental health nurses? An exploratory study of the formal written communication at the end of clinical nursing practice modules'. *Journal of Psychiatric and Mental Health Nursing,* 7 (5)*,* 407-416.

Buchel, T. L. and Edwards, F. D. (2005). 'Characteristics of effective clinical teachers'. *Family Medicine,* 37 (1)*,* 30-35.

Bussey, S. (2019). *Teaching undergraduate medical students: exploring the clinical teacher experience*. The Open University, Doctoral Dissertation.

Caballero, C. L., Walker, A. and Fuller-Tyszkiewicz, M. (2011). 'The work readiness scale (WRS): developing a measure to assess work readiness in college graduates'. *Journal of teaching and learning for graduate employability,* 2 (2)*,* 41-54.

Cant, S. and Sharma, U. (1999). *A new medical pluralism? Alternative medicine, doctors, patients and the state*. London: UCL Press.

Cant, S. L. and Sharma, U. (1996). 'Professionalization of complementary medicine in the United Kingdom'. *Complementary therapies in medicine,* 4 (3)*,* 157-162.

Cantillon, P., Dornan, T. and De Grave, W. (2019). 'Becoming a Clinical Teacher: Identity Formation in Context'. *Academic Medicine,* 94 (10)*,* 1610-1618.

Carr, W. (2005). 'The Role of Theory in the Professional Development of an Educational Theorist'. *Pedagogy, Culture and Society,* 13 (3).

Carraccio, C., Wolfsthal, S. D., Englander, R., Ferentz, K. and Martin, C. (2002). 'Shifting paradigms: from Flexner to competencies'. *Acad Med,* 77.

Chavez, C. (2008). 'Conceptualizing from the Inside: Advantages, Complications, and Demands on Insider Positionality'. *The Qualitative Report* 13 (3)*,* 472-494.

Cheek, C., Hays, R., Smith, J. and Allen, P. (2018). Improving case study research in medical education: a systematised review (Vol. 52, pp. 480-487).

Cleland, J. A., Knight, L. V., Rees, C. E., Tracey, S. and Bond, C. M. (2008). 'Is it me or is it them? Factors that influence the passing of underperforming students.'. *Medical education,* 42 (8)*,* 800-809.

CMBA. (2016). *Chinese medicine accreditation guidance materials Vs 3.0 Nov 2016*. [Online]. Available at: http://www.chinesemedicineboard.gov.au/Accreditation/Application-information.aspx. [Last accessed 20/01/2018].

CMBA. (2019). *Accreditation Standards: Chinese Medicine December*. [Online]. Available at: http://www.chinesemedicineboard.gov.au/Accreditation.aspx. [Last accessed 08/07/2020].

Collini, S. (2017). *Speaking of universities*. New York: Verso.

Cook, D., Dupras, D., Beckman, T., Thomas, K. and Pankratz, V. (2009). 'Effect of Rater Training on Reliability and Accuracy of Mini-CEX Scores: A Randomized, Controlled Trial'. *Journal of General Internal Medicine,* 24 (1)*,* 74-79.

Cooke, M., Irby, D. M. and O'Brien, B. C. (2010). *Educating physicians: a call for reform of medical school and residency*. CA: Jossey-Bass.

Creswell, J. W. (2007). *Qualitative inquiry & research design: choosing among five approaches*. (2nd Ed.). Thousand Oaks: Sage Publications

Crossley, J. (2013). 'Validity and truth in assessment'. *Medical Education,* 47 (12)*,* 1152-1154.

Crossley, J. (2016). 'Assessing the Non-Cognitive Domains: Measuring What Matters Well'. In E. Cano and G. Ion (Eds), *Innovative Practices for Higher Education Assessment and Measurement* (pp. 348-372): IGI Global.

Crossley, J., Johnson, G., Booth, J. and Wade, W. (2011). 'Good questions, good answers: Construct alignment improves the performance of workplace-based assessment scales'. *Medical Education,* 45 (6)*,* 560-569.

Crossley, J. and Jolly, B. (2012). 'Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people'. *Medical Education,* 46 (1)*,* 28-37.

Dahlberg, K. M. and Dahlberg, H. K. (2004). 'Description vs. interpretation - a new understanding of an old dilemma in human science research'. *Nurs Philos,* 5 (3)*,* 268-73.

Dall'Alba, G. and Sandberg, J. (2014). 'A phenomenological perspective on researching work and learning'. In S. Billett, C. Harteis and H. Gruber (Eds), *International Handbook of Research in Professional and Practice-based Learning* (pp. 279-304). Dordrecht: Springer.

Davey, N. (2006). *Unquiet understanding: Gadamer's philosophical hermeneutics.* Albany, NY: SUNY.

de Jonge, L. P., Timmerman, A. A., Govaerts, M. J., Muris, J. W., Muijtjens, A. M., Kramer, A. W. and van der Vleuten, C. P. (2017). 'Stakeholder perspectives on workplace-based performance assessment: towards a better understanding of assessor behaviour'. *Adv Health Sci Educ Theory Pract,* 22 (5)*,* 1213-1243.

Delandshere, G. and Petrosky, A. R. (1998). 'Assessment of Complex Performances: Limitations of Key Measurement Assumptions'. *Educational Researcher,* 27 (2)*,* 14-24.

Denscombe, M. (2010). *The Good Research Guide. For small-scale social research projects.* (4th Ed.). England: McGraw Hill, Open University Press.

Denzin, N. K. and Lincoln, Y. S. (2008). *The landscape of qualitative research.* (3rd Ed.). Los Angeles: Sage.

Docherty, M. (2017). *Situated competence: Towards a holistic definition of physician competence.* Unpublished Doctoral dissertation. Fielding Graduate University, Web.

Dornan, T. (2012). 'Review: Communication, knowing and being in work-based learning'. In V. Cook, C. Daly and M. Newman (Eds), *Work-based Learning in Clinical Settings.Insights from socio-cultural perspectives* (pp. 207-218). London: Radcliffe Publishing.

Dornan, T., Scherpbier, A., King, N. and Boshuizen, H. (2005). 'Clinical teachers and problem-based learning: a phenomenological study'. *Medical Education,* 39*,* 163-170.

Drake, P. and Heath, L. (2008). 'Insider research in schools and universities: the case of the professional doctorate'. In P. Sikes and A. Potts (Eds), *Researching Education from the Inside* (pp. 128-143). London: Routledge.

Dudek, N. L., Marks, M. B. and Regehr, G. (2005). 'Failure to fail: The perspective of clinical supervisors'. *Acad Med,* 80 (10)*,* S84–S87.

Duffy, R. (2013). 'Nurse to educator? Academic roles and the formation of personal academic identities'. *Nurse Education Today,* 33 (6)*,* 620-624.

Dunne, J. (2005). 'An Intricate Fabric: understanding the rationality of practice'. *Pedagogy, Culture and Society,* 13 (3).

Durning, S. J., Artino, A. R., Jr., Schuwirth, L. and van der Vleuten, C. (2013). 'Clarifying assumptions to enhance our understanding and assessment of clinical reasoning'. *Acad Med,* 88 (4)*,* 442-8.

Edwards, A. (2010). *Being an Expert Professional Practitioner. The Relational Turn in Expertise.* Dordrecht: Springer.

Eisner, E. W. (2004). 'What Can Education Learn from the Arts about the Practice of Education?'. *International Journal of Education & the Arts,* 5 (4)*,* 1-13.

Epstein, R. M. (2007). 'Assessment in medical education'. *N Engl J Med,* 356 (4)*,* 387-96.

Epstein, R. M. and Hundert, E. M. (2002). 'Defining and assessing professional competence'. *JAMA,* 287.

Eraut, M. (1994). *Developing professional knowledge and competence.* London: Falmer Press.

Eraut, M. (2000). 'Non-formal learning and tacit knowledge in professional work'. *British Journal Of Educational Psychology,* 70, 113-136.

Eraut, M. (2004). 'Informal learning in the workplace'. *Studies in Continuing Education,* 26 (2)*,* 247-273.

Eraut, M. (2010). 'Knowledge, Working Practices, and Learning'. In S. Billett (Ed.), *Learning through Practice. Models, Traditions, Orientations and Approaches* (pp. 37-58). Dordrecht: Springer.

Eva, K. W. and Hodges, B. D. (2012). 'Scylla or Charybdis? Can we navigate between objectification and judgement in assessment?'. *Medical Education,* 46 (9)*,* 914-9.

Evans, K. and Guile, D. (2012a). 'Putting different forms of knowledge to work in practice', *Practice-based education* (pp. 113-130): Brill Sense.

Evans, K. and Guile, D. (2012b). 'Putting different forms of knowledge to work in practice'. In J. Higgs, R. Barnett, S. Billett, M. Hutchins and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 113-130). Roterdam: Sense.

Evans, K., Guile, D. and Harris, J. (2011). 'Rethinking Work-Based Learning: For Education Professionals and Professionals Who Educate'. In M. Malloch, L. Cairns, K. Evans and B. O'Connor (Eds), *The SAGE Handbook of Workplace Learning* (pp. 149-161). Los Angeles, [Calif.]; London: Sage.

Farquhar, J. (1987). 'Problems of knowledge in contemporary Chinese medical discourse'. *Soc Sci Med,* 24 (12)*,* 1013-21.

Farquhar, J. (1994). *Knowing Practice. The Clinical Encounter of Chinese Medicine.* Boulder, CO: Westview Press.

Fenwick, T. (2012). 'Matterings of Knowing and Doing: Sociomaterial Approaches to Understanding Practice'. In P. Hager, A. Lee and A. Reich (Eds), *Practice, Learning and Change. Practice-Theory Perspectives on Professional Learning* (pp. 67-83). Dordrecht: Springer.

Fenwick, T. (2014). 'Sociomateriality in medical practice and learning: attuning to what matters'. *Medical Education,* 48 (1)*,* 44-52.

Fenwick, T., Nerland, M. and Jensen, K. (2012). 'Sociomaterial approaches to conceptualising professional learning and practice'. *Journal of Education and Work,* 25 (1)*,* 1-13.

Fenwick, T. J. and Nerland, M. (2014). 'Introduction: Sociomaterial professional knowing, work arrangements and responsibility: new times, new concepts?'. In T. J. Fenwick and M. Nerland (Eds), *Reconceptualising professional learning : sociomaterial knowledges, practices, and responsibilities.* New York: Routledge.

Fishbain, D., Danon, Y. and Nissanholz-Gannot, R. (2019). 'Accreditation systems for Postgraduate Medical Education: a comparison of five countries'. *Adv in Health Sci Educ,* 24 (3)*,* 503-524.

Flesch, H. (2013). 'A foot in both worlds: education and the transformation of Chinese medicine in the United States'. *Med Anthropol,* 32 (1)*,* 8-24.

Flick, U. (2004). 'Triangulation in qualitative research'. In U.Flick, E. v. Kardoff and I.Steinke (Eds), *A companion to qualitative research* (pp. 178-183). London: Sage.

Flyvbjerg, B. (2006). 'Five Misunderstandings About Case-Study Research'. *Qualitative Inquiry,* 12 (2)*,* 219-245.

Flyvbjerg, B. (2011). 'Case Study'. In N. K. Denzin and Y. S. Lincoln (Eds), *The Sage Handbook of Qualitative Research* (pp. 301-316). Thousand Oaks, California: Sage.

Frank, J. R., Mungroo, R., Ahmad, Y., Wang, M., De Rossi, S. and Horsley, T. (2010a). 'Toward a definition of competency-based education in medicine: a systematic review of published definitions'. *Medical Teacher,* 32 (8)*,* 631-637.

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., Harris, P., Glasgow, N. J., Campbell, C., Dath, D., Harden, R. M., Iobst, W., Long, D. M., Mungroo, R., Richardson, D. L., Sherbino, J., Silver, I., Taber, S., Talbot, M. and Harris, K. A. (2010b). 'Competency-based medical education: theory to practice'. *Med Teach,* 32 (8)*,* 638-45.

Freidson, E. (2001). *Professionalism: the third logic.* Chicago: University of Chicago Press.

Friedman, A. (2010). 'Perception: A cognitive sociological approach'. In F. Clement and L. Kaufman (Eds), *La Sociologie Cognitive.* Paris: Ophrys /Maison des Sciences de l'Homme, coll. Cogniprisme.

Gherardi, S. (2009a). 'Introduction: The critical power of the 'practice lens''. *Management Learning,* 40 (2)*,* 115-128.

Gherardi, S. (2009b). 'Knowing and learning in practice-based studies: an introduction'. *The Learning Organization,* 16 (5)*,* 352-359.

Gherardi, S. (2009c). 'Practice? It's a Matter of Taste!'. *Management Learning,* 40 (5)*,* 535-550.

Gherardi, S. (2019). 'Introduction'. In S. Gherardi (Ed.), *How to conduct a practice-based study : problems and methods* (2nd ed., pp. 1-7). Northampton, MA: Edward Elgar.

Gherardi, S., Nicolini, D. and Yanow, D. (2003). *Knowing in organizations : a practice-based approach.* London: M.E. Sharpe.

Gibson, W. and Brown, A. (2009). *Working with qualitative data.* London: Sage.

Gilbert, L. (2002). 'Going the distance: 'Closeness' in qualitative data analysis software'. *International Journal of Social Research Methodology,* 5 (3)*,* 215-228.

Gillies, D. (2010). *A brief critical dictionary of education.* [Online]. Available at: http://www.dictionaryofeducation.co.uk/component/glossary/Glossary-1/e/edumetric-300/. [Last accessed 08/07/2020].

Giloi, S. (2014). 'Design assessment: A socially responsible practice or subjective judgement?'. *Design with the Other,* 90*,* 234-240.

Gingerich, A., Kogan, J., Yeates, P., Govaerts, M. and Holmboe, E. (2014). 'Seeing the 'black box' differently: assessor cognition from three research perspectives'. *Medical education,* 48 (11)*,* 1055.

Gingerich, A., Regehr, G. and Eva, K. W. (2011). 'Rater-based assessments as social judgments: Rethinking the etiology of rater errors'. *Academic Medicine,* 86 (10)*,* S1-S7.

Ginsburg, S., McIlroy, J., Oulanova, O., Eva, K. and Regehr, G. (2010). 'Toward authentic clinical evaluation: pitfalls in the pursuit of competency'. *Acad Med,* 85 (5)*,* 780-6.

Ginsburg, S., Regehr, G., Lingard, L. and Eva, K. W. (2015). 'Reading between the lines: faculty interpretations of narrative evaluation comments'. *Medical Education,* 49 (3)*,* 296-306.

Givati, A. (2012). *The Holistic Discourse and Formalising Education of Non-medicallyQualified Acupuncturists and Homeopaths in England: an In-depth, Qualitative Study*. Unpublished PhD Thesis. University of Portsmouth, UK.

Givati, A. (2015). 'Performing 'pragmatic holism': Professionalisation and the holistic discourse of non-medically qualified acupuncturists and homeopaths in the United Kingdom'. *Health,* 19 (1)*,* 34-50.

Givati, A. and Hatton, K. (2015). 'Traditional acupuncturists and higher education in Britain: the dual, paradoxical impact of biomedical alignment on the holistic view'. *Soc Sci Med,* 131*,* 173-80.

Glatz, S. (2019). Elucidating the Relationship Between Chinese Medicine and Systems Biology : a Multi-Sited Ethnography: ProQuest Dissertations Publishing.

Gleeson, D. and Shain, F. (1999). 'Managing Ambiguity: Between Markets and Managerialism — A Case Study of 'Middle' Managers in Further Education'. *The Sociological review (Keele),* 47 (3)*,* 461-490.

GMC. (2018). *General Medical Council. Outcomes for Graduates*. [Online]. Available at: https://www.gmc-uk.org/education/standards-guidance-and-curricula/standards-and-outcomes/outcomes-for-graduates/outcomes-for-graduates. [Last accessed 12/05/2020].

Goffman, E. (1986). *Frame analysis: an essay on the organization of experience.* Boston: Northeastern University Press.

Goldszmidt, M. (2017). 'When I say … sociomateriality'. *Medical Education,* 51 (5)*,* 465-466.

Govaerts, M. and van der Vleuten, C. P. M. (2013). 'Validity in work‐based assessment: expanding our horizons'. *Medical Education,* 47 (12)*,* 1164-1174.

Govaerts, M. J., van de Wiel, M. W., Schuwirth, L. W., van der Vleuten, C. P. and Muijtjens, A. M. (2013). 'Workplace-based assessment: raters' performance theories and constructs'. *Adv Health Sci Educ Theory Pract,* 18 (3)*,* 375-96.

Govaerts, M. J., van der Vleuten, C. P., Schuwirth, L. W. and Muijtjens, A. M. (2007). 'Broadening perspectives on clinical performance assessment: rethinking the nature of in-training assessment'. *Adv Health Sci Educ Theory Pract,* 12 (2)*,* 239-60.

Govaerts, M. J. B., van der Vleuten, C. P. M. and Holmboe, E. S. (2019). 'Managing tensions in assessment: moving beyond either–or thinking'. *Medical Education,* 53 (1)*,* 64-75.

Grant, J. (2014). Principles of Curriculum Design. In T. Swanwick (ed), *Understanding Medical Education: Evidence, Theory and Practice* (2nd ed., pp. 31-46). Chichester: Wiley Blackwell.

Grealish, L. (2015). 'Professional Standards in Curriculum Design: A Socio-Technical Analysis of Nursing Competency Standards'. In M. Kennedy, S. Billett, S. Gherardi and L. Grealish (Eds), *Practice-based Learning in Higher Education: Jostling cultures* (pp. 85-98). Dordrecht: Springer.

Greene, M. J. (2014a). 'On the Inside Looking In: Methodological Insights and Challenges in Conducting Qualitative Insider Research'. *The Qualitative Report,* 19*,* 29.

Greene, M. J. (2014b). 'On the Inside Looking In: Methodological Insights and Challenges in Conducting Qualitative Insider Research'. *Qualitative Report,* 19 (29)*,* 1-13.

Grootenboer, P., Ronnerman, K. and Edwards-Groves, C. (2017). 'Leading from the Middle: A Praxis-Oriented Practice'. In P. Grootenboer, C. Edwards-Groves and S. Choy (Eds), *Practice Theory Perspectives on Pedagogy and Education* (pp. 243-263). Singapore: Springer.

Gruppen, D. L., ten Cate, A. O., Lingard, W. L., Teunissen, R. P. and Kogan, R. J. (2018). 'Enhanced Requirements for Assessment in a Competency-Based, Time-Variable Medical Education System'. *Academic Medicine,* 93 (Supplement 3S)*,* S17-S21.

Hager, P. (2004). 'The competence affair, or why vocational education and training urgently needs a new understanding of learning'. *Journal of Vocational Education & Training,* 56 (3)*,* 409-433.

Hager, P. (2011). 'Theories of workplace learning'. In L. C. M. Malloch, K. Evans and B. O'Connor (Ed.), *The SAGE Handbook of Workplace Learning* (pp. 17-31). Los Angeles: Sage.

Hager, P. (2012). 'Theories of Practice and Their Connections with Learning: A Continuum of More and Less Inclusive Accounts'. In P. Hager, A. Lee and A. Reich (Eds), *Practice, Learning and Change. Practice-Theory Perspectives on Professional learning* (pp. 17-32). Dordrecht: Springer.

Hager, P. (2017). 'The Integrated View on Competence'. In M. Mulder (Ed.), *Competence-based Vocational and Professional Education: Bridging the Worlds of Work and Education* (pp. 203-228). Cham: Springer International Publishing.

Hager, P. and Butler, J. (1996). 'Two models of educational assessment'. *Assessment and Evaluation in Higher Education,* 21 (4)*,* 367–78.

Hager, P. and Gonczi, A. (1996). 'What is competence?'. *Medical Teacher,* 18 (1)*,* 15-18.

Hager, P., Lee, A. and Reich, A. (2012). 'Problematising Practice, Reconceptualisating Learning and Imagining Change'. In P. Hager, A. Lee and A. Reich (Eds), *Practice, Learning and Change. Practice-Theory Perspectives on Professional learning* (pp. 1-14). Dordrecht: Springer.

Hamilton, D. (2005). 'Introduction: Knowing Practice'. *Pedagogy, Culture and Society,* 13 (3)*,* 285-289.

Harden, R. M. and Crosby, J. R. (1999). 'AMEE Guide No.14: Outcome-based education: Part 5--From competency to meta-competency: a'. *Medical Teacher,* 21 (6)*,* 546.

Harman, K. and McDowell, L. (2011). 'Assessment talk in Design: the multiple purposes of assessment in HE'. *Teaching in Higher Education,* 16 (1)*,* 41-52.

Hawe, E. (2003). ''It's Pretty Difficult to Fail': The reluctance of lecturers to award a failing grade'. *Assessment & Evaluation in Higher Education,* 28 (4)*,* 371-382.

Higgs, J. (2012a). 'Practice-based Education Pedagogy'. In J. Higgs, R. Barnett, S. Billett, M. Hutchings and F. Trede (Eds), *Practice-Based Education. Perspectives and Strategies* (pp. 71-80). Rotterdam: Sense Publishers.

Higgs, J. (2012b). 'Practice-based education. The practice-education-context-quality nexus'. In J. Higgs, R. Barnett, S. Billett, M. Hutchins and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 3-12). Rotterdam: Sense Publishers.

Higgs, J., Barnett, R., Billett, S., Hutchings, M. and Trede, F. (Eds) (2012), *Practice-Based Education. Perspectives and Strategies*. Rotterdam: Sense Publishers.

Higgs, J., Loftus, S. and Trede, F. (2010). 'Education for Future Practice'. In J. Higgs, D. Fish, I. Goulter, S. Loftus, J. Reid and F. Trede (Eds), *Education for Future Practice* (pp. 3-14). Rotterdam: Sense Publishers.

Ho, M. J., Lin, C. W., Chiu, Y. T., Lingard, L. and Ginsburg, S. (2012). 'A cross-cultural study of students' approaches to professional dilemmas: sticks or ripples'. *Medical Education,* 46 (3), 245-56.

Ho, M. J., Yu, K. H., Hirsh, D., Huang, T. S. and Yang, P. C. (2011). 'Does one size fit all? Building a framework for medical professionalism'. *Acad Med,* 86 (11), 1407-14.

Ho, M. J., Yu, K. H., Pan, H., Norris, J. L., Liang, Y. S., Li, J. N. and Hirsh, D. (2014). 'A tale of two cities: understanding the differences in medical professionalism between two Chinese cultural contexts'. *Acad Med,* 89 (6), 944-50.

Hodge, S., Mavin, T. and Kearns, S. (2019). 'Hermeneutic Dimensions of Competency-Based Education and Training'. *Vocations and Learning,* 13 (1), 27-46.

Hodges, B. (2013). 'Assessment in the post-psychometric era: Learning to love the subjective and collective'. *Medical Teacher,* 35 (7), 564-568.

Hodges, B. D. (2010). 'A tea-steeping or i-Doc model for medical education?'. *Acad Med,* 85 (9 Suppl), S34-44.

Hodges, B. D. (2012). 'The Shifting Discourses of Competence'. In B. D. Hodges and L. Lingard (Eds), *The Question of Competence. Reconsidering Medical Education in the Twenty-First Century* (pp. 14-41). Ithaca: Cornell University Press.

Hodges, B. D. and Lingard, L. (2012a). 'Introduction'. In B. D. Hodges and L. Lingard (Eds), *The Question of Competence. Reconsidering Medical Education in the Twenty-First Century* (pp. 1-13). Ithaca: Cornell University Press.

Hodges, B. D. and Lingard, L. (Eds) (2012b), *The Question of Competence. Reconsidering Medical Education in the Twenty-First Century*. Ithaca: Cornell University Press.

Hodwitz, K. R. (2018). *The influence of assessor training on scoring and feedback: A qualitative study of assessor perceptions*. University of Toronto (Canada), ProQuest Dissertations and Theses Global.

Holmboe, E. (2015). 'Direct observation of students' clinical skills'. In L. Pangaro and W. McGachie (Eds), *Handbook on medical student evaluation and assessment* (pp. 97-112). North Syracuse, New York: Gegensatz Press.

Holmboe, E. S., Edgar, L. and Hamstra, S. (2016). *The milestones guidebook*. Chicago, IL: Accreditation Council for Graduate Medical Education.

Holmboe, E. S., Sherbino, J., Englander, R., Snell, L., Frank, J. R. and ICBMECollaborators. (2017). 'A call to action: the controversy of and

rationale for competency-based medical education'. *Medical teacher,* 39 (6)*,* 574-581.

Holmboe, E. S., Sherbino, J., Long, D. M., Swing, S. R. and Frank, J. R. (2010). 'The role of assessment in competency-based medical education'. *Medical Teacher,* 32 (8)*,* 676-682.

Holstein, J. A. and Gubrium, J. F. (2004). 'The active interview'. In D. Silverman (Ed.), *Qualitative Research: Theory, Method and Practice* (Vol. 2, pp. 140-161). London: Sage.

Hopper, A. and Parrott, A. (2005). 'Reflective practice, professionalism and acupuncture education'. *European Journal of Oriental Medicine,* 5 (1)*,* 44-51.

Hopwood, N. (2016). 'Sociomaterialism, Practice Theory, and Workplace Learning'. In N. Hopwood (Ed.), *Professional Practice and Learning. Times, Spaces, Bodies, Things* (pp. 53-115). Switzerland: Springer.

Hordern, J. (2014). 'Productive Systems of Professional Formation'. In S. Billett, C. Harteis and H. Gruber (Eds), *International Handbook of Research in Professional and Practice-based Learning* (pp. 163-193). Dordrecht, Netherlands: Springer eBooks.

Hoyle, E. and Wallace, M. (2007). 'Educational Reform: An Ironic Perspective'. *Educational management, administration & leadership,* 35 (1)*,* 9-25.

Huddle, S. T. and Heudebert, R. G. (2007). 'Viewpoint: Taking Apart the Art: The Risk of Anatomizing Clinical Competence'. *Academic Medicine,* 82 (6)*,* 536-541.

Hughes, L. J., Mitchell, M. and Johnston, A. N. B. (2016). ''Failure to fail' in nursing – A catch phrase or a real issue? A systematic integrative literature review'. *Nurse Education in Practice,* 20*,* 54-63.

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)*,* 351-355.

Hurst, K. M. (2010). 'Experiences of new physiotherapy lecturers making the shift from clinical practice into academia'. *Physiotherapy,* 96 (3)*,* 240-247.

Hyland, T. (2017). 'Mindful Working and Skilful Means: Enhancing the Affective Elements of Vocational Education and Training Through the Ethical Foundations of Mindfulness'. In M. Mulder (Ed.), *Competence-based Vocational and Professional Education. Bridging the World of Work and Education* (pp. 145-164). Edinburgh: Springer.

ICBME. (2018). *International Competency-based Medical Education Collaborators.* [Online]. Available at: http://gocbme.org/icbme-site/competency-based-medical-education.html. [Last accessed 08/07/2020].

Innes, S. I., Cope, V., Leboeuf-Yde, C. and Walker, B. F. (2019). 'A perspective on Chiropractic Councils on Education accreditation standards and processes from the inside: A narrative description of expert opinion'. *Chiropractic and Manual Therapies,* 27 (1)*.*

Irby, M. D., Cooke, C. M. and O'Brien, C. B. (2010). 'Calls for Reform of Medical Education by the Carnegie Foundation for the Advancement of Teaching: 1910 and 2010'. *Academic Medicine,* 85 (2)*,* 220-227.

Jawitz, J. (2009). 'Learning in the academic workplace: The harmonization of the collective and individual habitus'. *Studies in Higher Education,* 34 (6)*,* 601-614.

Jessup, G. (1991). *Outcomes : NVQs and the Emerging Model of Education and Training*. London: The Falmer Press.

Kemmis, R. B. (2008). 'Freedom for praxis: An unburied and unforgotten tradition'. In S. Kemmis and T. Smith (Eds), *Enabling Praxis: Challenges for Education* (pp. 195-215). Rotterdam:: Sense Publishers.

Kemmis, S. (2009). 'Understanding Professional Practice: A Synoptic Framework'. In B. Green (Ed.), *Understanding and Researching Professional Practice* (pp. 19-38). Rotterdam: Sense Publishers.

Kemmis, S. (2012). 'Pedagogy, praxis and practice- based higher education'. In J. Higgs, R. Barnett, S. Billett, M. Hutchins and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 81-100). Roterdam: Sense.

Kemmis, S., Edwards-Groves, C., Wilkinson, J. and Hardy, I. (2012). 'Ecologies of Practices'. In P. Hager, A. Lee and A. Reich (Eds), *Practice, Learning and Change: Practice-Theory Perspectives on Professional Learning* (pp. 33-49). Dordrecht: Springer Netherlands.

Kemmis, S. and Grootenboer, P. (2008). 'Situating praxis in Practice: Practice architectures and the cultural, social and material conditions for practice'. In S. Kemmis and T. Smith (Eds), *Enabling praxis: Challenges for education* (pp. 37-62). Rotterdam: Sense Publishers.

Kemmis, S. and Mahon, K. (2017). 'Practice architectures of university education'. In P. Grootenboer, C. Edwards-Groves and S. Choy (Eds), *Practice Theory Perspectives on Pedagogy and Education. Praxis, Diversity and Contestation* (pp. 107-141). Singapore: Springer.

Kemmis, S., Wilkinson, J., Edwards-Groves, C., Hardy, I., Grootenboer, P. and Bristol, L. (2014). *Changing practices, changing education*. Singapore: Springer Science & Business Media.

Kennedy, C. (2014). *Update on Outcome/Competency-based Medical Education*. [Online]. Available at: https://amee.org/getattachment/Marketing-Toolkit/38074-Outcome_Competency-Med-Ed-WEB.PDF. [Last accessed 08/07/2020].

Kennedy, M., Billett, S., Gherardi, S. and Grealish, L. (2015). 'Practice-based learning in higher education: Jostling cultures'. In M. Kennedy, S. Billett, S. Gherardi and L. Grealish (Eds), *Practice-based Learning in Higher Education: Jostling cultures* (pp. 1-13). Dordrecht: Springer.

Kilgour, A., Gerzina, T., Keppell, M. and Gerzina, J. (2013). 'Understanding the Place of Assessment Standards. A Case Study in Medical Diagnostic Radiography'. In M. Smith and T. Levett-Jones (Eds), *Providing clinical education: Working across sectors.* (pp. 277-286). Rotterdam: Sense Publishers.

Kilminster, S. M. and Jolly, B. C. (2000). 'Effective supervision in clinical practice settings: a literature review'. *Medical Education,* 34 (10)*,* 827-840.

Klamen, D. L., Williams, R. G., Roberts, N. and Cianciolo, A. T. (2016). 'Competencies, milestones, and EPAs – Are those who ignore the past condemned to repeat it?'. *Medical Teacher,* 38 (9)*,* 904-910.

Knight, P. (2007). 'Grading, classifying and future learning'. In D. Boud and N. Falchikov (Eds), *Rethinking Assessment in higher Education. Learning for the Longer Term* (pp. 72-86). Oxon: Routledge.

Kogan, J. R., Conforti, L., Bernabeo, E., Iobst, W. and Holmboe, E. (2011). 'Opening the black box of clinical skills assessment via observation: a conceptual model'. *Med Educ,* 45 (10)*,* 1048-60.

Kogan, J. R., Conforti, L. N., Bernabeo, E., Iobst, W. and Holmboe, E. (2015). 'How faculty members experience workplace‐based assessment rater training: a qualitative study'. *Medical education,* 49 (7)*,* 692-708.

Kogan, J. R. and Holmboe, E. (2013). 'Realizing the promise and importance of performance-based assessment'. *Teach Learn Med,* 25 Suppl 1*,* S68-74.

Kogan, J. R. and Holmboe, E. S. (2018). 'Direct Observation'. In E. S. Holmboe, S. J. Durning and R. E. Hawkins (Eds), *Practical Guide to the Evaluation of Clinical Competence* (pp. 323-512). Philadelphia: Elsevier.

Kuper, A., Whitehead, C. and Hodges, B. D. (2013). 'Looking back to move forward: using history, discourse and text in medical education research: AMEE guide no. 73'. *Med Teach,* 35 (1)*,* e849-60.

Kvale, S. (1994). 'Ten Standard Objections to Qualitative Research Interviews'. *Journal of Phenomenological Psychology,* 25 (2)*,* 147-173.

Lammi, I. J. (2018). *A Practice Theory in Practice. Analytical Consequences in the Study of Organization and Socio-Technical Change.* Unpublished PhD. Uppsala University.

Landri, P. (2012). 'A Return to Practice: Practice-Based Studies of Education'. In P. Hager, A. Lee and A. Reich (Eds), *Practice, Learning and Change. Practice-Theory Perspectives on Professional Learning* (pp. 85-100). Dordrecht: Springer.

Lei, S. H.-l. (2002). 'How Did Chinese Medicine Become Experiential? The Political Epistemology of Jingyan'. *Positions,* 10 (2)*,* 333-364.

Lillrank, A. (2012). Managing the Interviewer Self. In J. F. Gubrium, J. A. Holstein, A. B. Marvasti and K. D. McKinney (eds), *The SAGE Handbook of Interview Research: The Complexity of the Craft* (pp. 281-294). Thousand Oaks, Calif.: Sage.

Lingard, L. (2012). 'Rethinking Competence in the Context of Teamwork'. In B. Hodges and L. Lingard (Eds), *The Question of Competence. Reconsidering Medicial Education int eh Twenty-First Century* (pp. 42-69). Ithaca: Cornell University Press.

Lockyer, J., Carraccio, C., Chan, M. K., Hart, D., Smee, S., Touchie, C., Holmboe, E. S., Frank, J. R. and Collaborators, I. (2017). 'Core principles of assessment in competency-based medical education'. *Med Teach,* 39 (6)*,* 609-616.

Lucey, C. R., Thibault, G. E. and ten Cate, O. (2018). 'Competency-Based, Time-Variable Education in the Health Professions: Crossroads'. *Academic Medicine,* 93 (3S)*,* S1-S5.

Lurie, S. J., Mooney, C. J. and Lyness, J. M. (2009). 'Measurement of the general competencies of the accreditation council for graduate medical education: a systematic review'. *Acad Med,* 84 (3)*,* 301-9.

Lynch, K. (2014). 'Control by numbers: new managerialism and ranking in higher education'. *Critical Studies in Education,* 56 (2)*,* 190-207.

MacLeod, A. (2011). 'Caring, competence and professional identities in medical education'. *Adv Health Sci Educ Theory Pract,* 16 (3)*,* 375-94.

Mahon, K., Kemmis, S., Francisco, S. and Lloyd, A. (2017). 'Introduction: Practice Theory and the Theory of Practice Architectures'. In K. Mahon, S. Francisco and S. Kemmis (Eds), *Exploring Education and Professional Practice. Through the Lens of Practice Architectures* (pp. 1-30). Singapore: Springer Singapore.

McColl, T. (2017). *MEdIC Series: Case of the Failure to Fail–Expert Review and Curated Community Commentary*. [Online]. Available at: https://www.aliem.com/2017/06/medic-case-failure-to-fail-expert-review-curated-community-commentary/. [Last accessed 08/07/2020].

McEwen, C. and Trede, F. (2014). 'The Academisation of Emerging Professions: Implications for Universities, Academics and Students'. *Power and Education,* 6 (2)*,* 145-154.

McGaghie, W. C., Miller, G. E., Sajid, A. W. and Telder, T. V. (1978). *Competency-Based Curriculum Development in Medical Education. An Introduction*. Public Health Papers No. 68.

McNamara, J. (2013). 'The challenge of assessing professional competence in work integrated learning'. *Assessment & Evaluation in Higher Education,* 38 (2)*,* 183-197.

Meddings, F. S. (2017). *Investigating assessment in Higher Education: demystifying marking and grading to reveal expert practices. A phenomenological analysis of marking and grading practices of novice and experienced health academics*. Unpublished PhD. University of Bradford, Faculty of Health Studies.

Mejicano, G. C. and Bumsted, T. N. (2018). 'Describing the Journey and Lessons Learned. Implementing a Competency-Based, Time-Variable Undergraduate Medical Education Curriculum'. *Acad Med,* 93 (3S)*,* S42-S48.

Merleau-Ponty, M. (2013). *Phenomenology of Perception*. (D. A. Landes, Trans.). London & New York: Routledge.

Miles, M. B., Huberman, A. M. and Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook*. (3rd Ed.). Thousand Oaks, California: Sage.

Moir, F. (2003). 'An exploration of clinical teaching using a reflective method with adult students'. *J Altern Complement Med,* 9 (5)*,* 789-93.

Moir, F. (2005). 'A day in the life of a reflective teacher: a reflective approach to research in education'. *European Journal of Oriental Medicine,* 5 (1)*,* 13-23.

Moir, F. (2014). An exploration of how practitioners make meaning in a Chinese medicine consultation and treatment. London: Institute of Education (unpublished IFS).

Morcke, A. M., Dornan, T. and Eika, B. (2013). 'Outcome (competency) based education: an exploration of its origins, theoretical basis, and empirical evidence'. *Advances in Health Sciences Education,* 18 (4)*,* 851-863.

Morris, C. (2019). 'Work-based Learning'. In T. Swanwick, K. Forrest and B. O'Brien (Eds), *Understanding Medical Education: Evidence, Theory and Practice* (3rd ed ed., pp. 940-1030). Hoboken, NJ: Wiley Blackwell

Natesan, P., Batley, N. J., Bakhti, R. and El-Doueihi, P. Z. (2018). 'Challenges in measuring ACGME competencies: considerations for milestones'. *International Journal of Emergency Medicine,* 11 (1)*,* 39.

Norman, G., Norcini, J. and Bordage, G. (2014). 'Competency-based education: milestones or millstones?'. *Journal of graduate medical education,* 6 (1)*,* 1.

O'Neal, C. (2016). 'Beyond Authenticity: What Should We Value in Assessment in Professional Education'. In P. F. Wimmers and M. Mentkowski (Eds), *Assessing Competence in Professional Performance Across Disciplines and Professions* (Vol. 13, pp. 51-71). Switzerland: Springer.

O'Hagan, S. R. and Wigglesworth, G. (2014). 'Who's Marking My Essay? The Assessment of Non-native-speaker and Native-speaker Undergraduate

Essays in an Australian Higher Education Context'. *Studies in Higher Education,* 14 (9), 1729-1747.

Orr, K., Cunliffe, A. L. and Bennett, M. (2009). 'Reflexivity in the co‐production of academic‐practitioner research'. *Qualitative Research in Organizations and Management: An International Journal,* 4 (1), 85-102.

Orr, S. (2007). 'Assessment moderation: constructing the marks and constructing the students'. *Assessment & Evaluation in Higher Education,* 32 (6), 645-656.

Orr, S. and Bloxham, S. (2013). 'Making judgements about students making work: Lecturers' assessment practices in art and design'. *Arts and Humanities in Higher Education,* 12 (2-3), 234-253.

Orrell, J. and Higgs, J. (2012). 'Social and political change: Implications for professional and practice-based university education'. In J. Higgs, R. Barnett, S. Billett, M. Hutchings and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 41-54). Rotterdam: Sense.

Pangaro, L. and ten Cate, O. (2013). 'Frameworks for learner assessment in medicine: AMEE Guide No. 78'. *Med Teach,* 35 (6), 197-210.

Park, Y. S., Hodges, B. D. and Tekian, A. (2016). 'Evaluating the Paradigm Shift from Time-Based Toward Competency-Based Medical Education: Implications for Curriculum and Assessment'. In P. F. Wimmers and M. Mentkowski (Eds), *Assessing Competence in Professional Performance across Disciplines and Professions* (pp. 411-425). Switzerland: Springer International Publishing.

Polit, D. F. and Beck, C. T. (2010). 'Generalization in quantitative and qualitative research: Myths and strategies'. *International Journal of Nursing Studies,* 47 (11), 1451-1458.

Preston, J. (2017). *Competence Based Education and Training (CBET) and the End of Human Learning.* Switzerland: Palgrave Macmillan.

Price, M., Carroll, J., O'Donovan, B. and Rust, C. (2011). 'If I was going there I wouldn't start from here: a critical commentary on current assessment practice'. *Assessment & Evaluation in Higher Education,* 36 (4), 479-492.

Prior, L. (2003). *Using documents in social research.* London: Sage Publications.

Proitz, T. S. (2010). 'Learning Outcomes: What Are They? Who Defines Them? When and Where Are They Defined?'. *Educational Assessment, Evaluation and Accountability,* 22 (2), 119-137.

Reckwitz, A. (2016). 'Practices and their affects'. In A. Hui, T. Schatzki and E. Shove (Eds), *The Nexus of Practices: Connections, Constellations and Practitioners.* London: Routledge.

Rees, C. E. (2004). 'The problem with outcomes-based curricula in medical education: insights from educational theory'. *Medical Education,* 38 (6), 593-8.

Regehr, G., Bogo, M., Regehr, C. and Power, R. (2007). 'Can we build a better mousetrap? Improving the measures of practice performance in the field practicum'. *Journal of Social Work Education,* 43 (2), 327-344.

Reich, A. and Hager, P. (2014). 'Problematising practice, learning and change: practice-theory perspectives on professional learning'. *Journal of Workplace Learning,* 26 (6/7), 418-431.

Rencic, J., Durning, S. J., Holmboe, E. and Gruppen, L. D. (2016). 'Understanding the Assessment of Clinical Reasoning'. In P. F. Wimmers and M. Mentkowski

(Eds), *Assessing Competence in Professional Performance across Disciplines and Professions* (pp. 209-235). Cham: Springer International Publishing.

Rethans, J. J., Norcini, J. J., Baron-Maldonado, M., Blackmore, D., Jolly, B. C., LaDuca, T., Lew, S., Page, G. G. and Southgate, L. H. (2002). 'The relationship between competence and performance: implications for assessing practice performance'. *Med Educ,* 36 (10)*,* 901-9.

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)*,* 80-86.

Robson, C. (2011). *Real World Research : A resource for users of social research methods in applied settings.* (3rd Ed.). Chichester: John Wiley & Sons.

Rošker, J. (2018). 'Epistemology in Chinese Philosophy'. [Online]. *The Stanford Encyclopedia of Philosophy* Available at: https://plato.stanford.edu/archives/fall2018/entries/chinese-epistemology. [Last accessed 26/03/2020].

Rotthoff, T. (2018). 'Standing up for Subjectivity in the Assessment of Competencies'. *GMS J Med Educ,* 35 (3)*,* Doc29.

Rust, C. (2007). 'Towards a scholarship of assessment'. *Assessment & Evaluation in Higher Education,* 32 (2)*,* 229-237.

Sadler, D. R. (2009). 'Indeterminacy in the use of preset criteria for assessment and grading'. *Assessment & Evaluation in Higher Education,* 34 (2)*,* 159-179.

Sadler, D. R. (2010). 'Fidelity as a precondition for integrity in grading academic achievement'. *Assessment & Evaluation in Higher Education,* 35 (6)*,* 727-743.

Saks, M. (2001). 'Alternative Medicine and the Health Care Division of Labour: Present Trends and Future Prospects'. *Current sociology,* 49 (3)*,* 119-134.

Saks, M. (2015). 'Power and Professionalisation in CAM f'. In N. Gale and J. V. McHale (Eds), *Handbook of complementary and alternative medicine. Perspectives from social science and law* (pp. 30-40). Abingdon, Oxon: Routledge.

Sambell, K., McDowell, L. and Montgomery, C. (2012). *Assessment for learning in higher education*. London: Routledge.

Sandelowski, M. and Leeman, J. (2012). 'Writing usable qualitative health research findings'. *Qual Health Res,* 22 (10)*,* 1404-13.

Schatzki, T. (1997). 'Practices and actions - A Wittgensteinian critique of Bourdieu and Giddens'. *Philosophy Of The Social Sciences,* 27 (3)*,* 283-308.

Schatzki, T. (2001). 'Introduction: practice theory'. In T. Schatzki, K. Knorr Cetina and E. von Savigny (Eds), *The Practice Turn in Contemporary Theory* (pp. 10-23). London: Routledge.

Schatzki, T. (2002). *The Site of the Social: A Philosophical Account of the Constitution of Social Life and Change*. University Park, PA: Pennsylvania State University Press.

Schatzki, T. (2012). 'A primer on practices: Theory and Research'. In J. Higgs, R. Barnett, S. Billett, M. Hutchins and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 13-26). Roterdam: Sense.

Schatzki, T. (2016). 'Sayings, texts and discursive formations'. In A. Hui, T. Schatzki and E. Shove (Eds), *The Nexus of Practices: Connections, Constellations, Practitioners* (pp. 415-). London: Routledge.

Schatzki, T. (2017). 'Practices and Learning'. In P. Grootenboer, C. Edwards-Groves and S. Choy (Eds), *Practice Theory Perspectives on Pedagogy and Education* (pp. 23-44). Singapore: Springer.

Schatzki, T., Knorr Cetina, K. and von Savigny, E. (2001). *The Practice Turn in Contemporary Theory*. London: Routledge.

Schatzki, T. R. (2011). 'Where the action is (on large social phenomena such as sociotechnical regimes)'. *Sustainable Practices Research Group, Working Paper,* 1*,* 1-31.

Scheid, V. (2002). *Chinese Medicine in Contemporary China: Plurality and Synthesis*. Durham & London: Duke University Press.

Scheid, V. (2008). 'Globalising Chinese Medical Understandings of Menopause'. *East Asia Science, Technology and Society: an International Journal,* 2*,* 485-506.

Scheid, V. (2016). 'Holism, Chinese Medicine and Systems Ideologies: Rewriting the Past to Imagine the Future'. In J. Richards, S. Atkinson, J. Macnaughton, A. Woods and A. Whitehead (Eds), *The Edinburgh Companion to the Critical Medical Humanities* (pp. 66-86). Edinburgh: Edinburgh University Press.

Scheid, V. and Karchmer, E. (2016). 'History of Chinese medicine, 1890 - 2010'. In V. Goosaert, J. Kiely and J. Lagerwey (Eds), *Modern Chinese Religion II: 1850-2015* (Vol. 1, pp. 141-196.). Leiden/Boston: Brill.

Scheid, V. and Lei, S. H.-l. (2014). 'The Institutionalization of Chinese Medicine'. In B. Andrews and M. B. Bullock (Eds), *Medical Transitions in Twentieth-Century China* (pp. 244-266). Bloomington and Indianapolis: Indiana University Press.

Schubert, W. H. (2008). 'Curriculum Inquiry'. In F. M. Connelly, M. F. He and J. Phillion (Eds), *The SAGE handbook of curriculum and instruction* (pp. 399-419). London: SAGE.

Schuwirth, L. and Van der Vleuten, C. (2010). 'Assessment in problem-based learning'. In H. van Berkel, A. Scherpbier, H. Hillen and C. P. M. van der Vleuten (Eds), *Lessons from Problem-based Learning* (pp. 193-202). New York: Oxford University Press.

Schuwirth, L. and van der Vleuten, C. (2012). 'Assessing Competence. Extending the Approaches to Reliability'. In B. D. Hodges and L. Lingard (Eds), *The Question of Competence. Reconsidering Medical Education in the Twenty-First Century* (pp. 113-130). Ithaca: Cornell University Press.

Schuwirth, L. and van der Vleuten, C. (2019). How to Design a Useful Test: The Principles of Assessment. In T. Swanwick, K. Forrest and B. O'Brien (eds), *Understanding Medical Education: Evidence, Theory and Practice* (3rd ed., pp. 1569-1650). Hoboken, NJ.: Wiley Blackwell

Schuwirth, L. W. and van der Vleuten, C. P. (2006). 'A plea for new psychometric models in educational assessment'. *Medical Education,* 40 (4)*,* 296-300.

Schuwirth, L. W. and Van der Vleuten, C. P. (2011a). 'Programmatic assessment: from assessment of learning to assessment for learning'. *Medical Teacher,* 33 (6)*,* 478-485.

Schuwirth, L. W. T. and Van der Vleuten, C. P. M. (2011b). 'Programmatic assessment: from assessment of learning to assessment for learning'. *Med Teach,* 33.

Schuwirth, L. W. T. and van der Vleuten, C. P. M. (2014). 'How to Design a Useful Test: The Principles of Assessment'. In T. Swanwick (Ed.), *Understanding*

Medical Education: Evidence, Theory and Practice (2 ed., pp. 241-254). Chichester: Wiley Blackwell.

Schwandt, T. (2005). 'On Modeling Our Understanding of the Practice Fields'. *Pedagogy, Culture and Society,* 13 (3)*,* 313-332.

Sebok‑Syer, S. S., Chahine, S., Watling, C. J., Goldszmidt, M., Cristancho, S. and Lingard, L. (2018). 'Considering the interdependence of clinical performance: implications for assessment and entrustment'. *Medical Education,* 52 (9)*,* 970-980.

Sfard, A. (1998). 'On Two Metaphors for Learning and the Dangers of Choosing Just One'. *Educational Researcher,* 27 (2)*,* 4-13.

Shanahan, E. M., van der Vleuten, C. and Schuwirth, L. (2019). 'Conflict between clinician teachers and their students: the clinician perspective'. *Advances in health sciences education: theory and practice.*

Shay, S. (2005). 'The assessment of complex tasks: a double reading'. *Studies in Higher Education,* 30 (6)*,* 663-679.

Shay, S. B. (2004). 'The Assessment of Complex Performance: A Socially Situated Interpretive Act'. *Harvard Educational Review,* 74 (3)*,* 307-329.

Sherbino, J., Kulasegaram, K., Worster, A. and Norman, G. R. (2013). 'The reliability of encounter cards to assess the CanMEDS roles'. *Advances in Health Sciences Education,* 18 (5)*,* 987-996.

Shorey, S., Lau, T. C., Lau, S. T. and Ang, E. (2019). 'Entrustable professional activities in health care education: a scoping review'. *Med Educ,* 53 (8)*,* 766-777.

Shulman, L. S. (2005). 'Signature pedagogies in the professions.'. *Daedalus,* 134 (3)*,* 52-59.

Shumway, J. M. and Harden, R. M. (2003). 'AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician'. *Med Teach,* 25 (6)*,* 569-84.

Sikes, P. and Potts, A. (2008). 'Closing thoughts'. In P. Sikes and A. Potts (Eds), *Researching Education from the Inside* (pp. 177-180). London: Routledge.

Silverman, D. (2017). *Doing Qualitative Research.* (5th Ed.). Los Angeles: Sage.

Smith, Sr. and Dollase, R. (1999). 'AMEE guide No. 14: Outcome-based education: Part 2 - Planning, implementing and evaluating a competency-based curriculum'. *Medical Teacher,* 21 (1)*,* 15-22.

Smith, M. and Levett-Jones, T. (2013). *Providing clinical education: Working across sectors.* Rotterdam: Sense Publishers.

Stake, R. E. (1995). *The art of case study research.* Thousand Oaks, California: Sage.

Stake, R. E. (2005). 'Qualitative case studies'. In N. K. Denzin and Y. S. Lincoln (Eds), *The Sage Handbook of Qualitative Research* (3rd ed., pp. 443-466). Thousand Oaks, CA: Sage.

Steinert, S. Y., O'Sullivan, M. P. and Irby, M. D. (2019). 'Strengthening Teachers' Professional Identities Through Faculty Development'. *Academic Medicine,* 94 (7)*,* 963-968.

Stone, C., Boud, D. and Hager, P. (2011). 'Assessment of osteopaths: Developing a capability-based approach to reviewing readiness to practice.'. *International Journal of Osteopathic Medicine,* 14 (4)*,* 129-140.

Talbot, M. (2004). 'Monkey see, monkey do: A critique of the competency model in graduate medical education'. *Medical Education,* 38 (6)*,* 587-592.

Tavares, W. (2014). *Impact of Mental Workload on Rater Performance and Behaviour*. Unpublished PhD. McMaster, Hamilton, Ontario.

Tavares, W., Kuper, A., Kulasegaram, K. and Whitehead, C. (2019). 'The compatibility principle: on philosophies in the assessment of clinical competence'. *Advances in Health Sciences Education,* Nov (1)*,* 1-16.

Tavares, W., Rowland, P., Dagnone, D., McEwen, L. A., Billett, S. and Sibbald, M. (2020). 'Translating outcome frameworks to assessment programmes: Implications for validity'. *Medical education*.

Taylor, K. (2004a). *Chinese medicine in early communist China, 1945-1963 : Medicine of revolution*. London: Routledge.

Taylor, K. (2004b). 'Divergent Interests and cultivated Misunderstandings: The Influence of the West on Modern Chinese Medicine'. *Social History of Medicine,* 17 (1)*,* 93-111.

Tekian, A., Hodges, B. D., Roberts, T. E., Schuwirth, L. and Norcini, J. (2015). 'Assessing competencies using milestones along the way'. *Medical Teacher,* 37 (4)*,* 399-402.

Tekian, A. and Norcini, J. J. (2016). 'Faculty Development in Assessment: What the Faculty Need to Know and Do'. In P. F. Wimmers and M. Mentkowski (Eds), *Assessing Competence in Professional Performance across Disciplines and Professions* (pp. 355-374). Switzerland: Springer International Publishing.

ten Cate, O. (2017). 'Competency-Based Postgraduate Medical Education: Past, Present and Future'. *GMS J Med Educ,* 34 (5)*,* Doc69.

ten Cate, O. and Billett, S. (2014). 'Competency-based medical education: origins, perspectives and potentialities'. *Med Educ,* 48 (3)*,* 325-32.

ten Cate, O. and Scheele, F. (2007). 'Competency-based postgraduate training: can we bridge the gap between theory and clinical practice?'. *Acad Med,* 82 (6)*,* 542–547.

Trede, F. and Smith, M. (2012). 'Challenges of assessment in practice-based education'. In J. Higgs, R. Barnett, S. Billett, M. Hutchings and F. Trede (Eds), *Practice - based Education. Perspectives and Strategies* (pp. 187-198). Roterdam: Sense.

Trede, F. and Smith, M. (2014). 'Workplace educators' interpretations of their assessment practices: a view through a critical practice lens'. *Assessment & Evaluation in Higher Education,* 39 (2)*,* 154-167.

Trowler, P. (2001). 'Captured by the discourse? The socially constitutive power of new higher education discourse in the UK.'. *Organization,* 8 (2)*,* 183-201.

Tummons, J. (2020). 'Education as a mode of existence: A Latourian inquiry into assessment validity in higher education'. *Educational Philosophy and Theory,* 52 (1)*,* 45-54.

Uddin, J. (2008). 'Chronological History of Chinese Medicine Education in the UK'. In M. McCarthy, S. Birch and A.-M. Findgott (Eds), *Thieme Almanac. Acupuncture and Chinese Medicine* (pp. 242). Stuttgart, Germany: Thieme.

van der Vleuten, C. P. (1996). 'The assessment of professional competence: Developments, research and practical implications'. *Adv Health Sci Educ Theory Pract,* 1 (1)*,* 41-67.

van der Vleuten, C. P., Schuwirth, L. W., Scheele, F., Driessen, E. W. and Hodges, B. (2010). 'The assessment of professional competence: building blocks for theory development'. *Best Pract Res Clin Obstet Gynaecol,* 24 (6)*,* 703-19.

van Lankveld, T., Schoonenboom, J., Croiset, G., Volman, M. and Beishuizen, J. (2017). 'The role of teaching courses and teacher communities in strengthening the identity and agency of teachers at university medical centres'. *Teaching and Teacher Education,* 67, 399-409.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C. and Herrera-Seda, C. (2018). 'Authentic assessment: creating a blueprint for course design'. *Assessment & Evaluation in Higher Education,* 43 (5), 840-854.

Wass, V. and Archer, J. (2011). 'Assessing learners'. In T. Dornan, K. Mann, A. Scherpbier and J. Spencer (Eds), *Medical Education. Theory and Practice* (pp. 229-255). Edinburgh: Elsevier.

Watling, C. J., Ajjawi, R. and Bearman, M. (2020). 'Approaching culture in medical education: Three perspectives'. *Medical Education,* 54 (4), 289-295.

Wear, D. (2008). 'On Outcomes and Humility'. *Academic Medicine,* 83 (7), 625-626.

Wengraf, T. (2001). *Qualitative research interviewing: Biographic narrative and semi-structured methods*. London: Sage.

Wheelahan, L. (2009). 'The Problem with CBT (and Why Constructivism Makes Things Worse)'. *Journal of Education and Work,* 22 (3), 227-242.

White, C. B. and Fantone, J. C. (2010). 'Pass–fail grading: laying the foundation for self-regulated learning'. *Advances in Health Sciences Education,* 15 (4), 469-477.

Whitehead, C. R., Kuper, A., Hodges, B. and Ellaway, R. (2015). 'Conceptual and practical challenges in the assessment of physician competencies'. *Med Teach,* 37 (3), 245-51.

Whitty, G. (2008). 'Changing modes of teaching professionalism: traditional, managerial, collaborative and democratic.'. In B. Cunningham (Ed.), *Exploring Professionalism.* (pp. 28-49). London: Institute of Education.

Wilkerson, L. and Doyle, L. H. (2011). 'Developing teachers and developing learners'. In T. Dornan, K. Mann, A. Scherpbier and J. Spencer (Eds), *Medical Education. Theory and Practice* (pp. 317-338). Edinburgh: Elsevier.

Williams, R. G., Klamen, D. A. and McGaghie, W. C. (2003). 'Cognitive, social and environmental sources of bias in clinical performance ratings'. *Teach Learn Med,* 15 (4), 270-92.

Winstone, N. E. and Boud, D. (2020). 'The need to disentangle assessment and feedback in higher education'. *Studies in higher education (Dorchester-on-Thames),* 1-12.

Wu, W.-Y., Yang, W.-Z., Hou, J.-J. and Guo, D.-A. (2015). 'Current Status and Future Perspective in the Globalization Of Traditional Chinese Medicines'. *World Journal of Traditional Chinese Medicine,* 1 (1), 1-4.

Yepes-Rios, M., Dudek, N., Duboyce, R., Curtis, J., Allard, R. J. and Varpio, L. (2016). 'The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42'. *Medical Teacher,* 38 (11), 1092-1099.

Yin, R. (2014). *Case study research: design and methods*. (5th Ed.). Los Angeles; London: Sage.

Yorke, M. (2005). *Issues in the assessment of practice-based professional learning. A Report prepared for the Practice-based Professional Learning CETL at the Open University*. Milton Keynes: Open University.

Yorke, M. (2009). 'Faulty Signals? Inadequacies of Grading Systems and a Possible Response.'. In G. Joughin (Ed.), *Assessment, Learning and Judgement in Higher Education* (pp. 65-84). B.V.: Springer.

Yorke, M. (2011a). 'Summative assessment: dealing with the 'measurement fallacy''. *Studies in Higher Education,* 36 (3)*,* 251-273.

Yorke, M. (2011b). 'Work-Engaged Learning: Towards a Paradigm Shift in Assessment Quality in Higher Education'. *Quality in Higher Education,* 17 (1)*,* 117–130.

Young, M. (2008). *Bringing Knowledge Back In.* London: Routledge.

Young, M. (2013). 'Overcoming the crisis in curriculum theory: a knowledge-based approach'. *Journal Of Curriculum Studies,* 45 (2)*,* 101-118.

Zhan, M. (2009). *Other-Wordly. Making Chinese Medicine through Transnational Frames* Durham: Duke University Press.

****

Appendices

Appendix 1   Manager Information Statement

An Exploration of the Assessment of Professional Practice in Chinese Medicine Institutions in the West

---

You are invited to participate in a research project being conducted by Felicity Moir as part of her Education Doctorate at the Institute of Education, London, UK.   Felicity is a Principal lecturer and Course Leader of a BSc (Hons) Chinese Medicine: Acupuncture and MSc Chinese Herbal Medicine professional entry courses at the University of Westminster, London, UK.

1. **What is the study about?**
   This study hopes to explore the way that Chinese medicine (CM) Higher Education institutions in the West assess the professional practice of their students. While the assessment of professional practice in the biomedical health professions has been much debated and researched, for example, Billett (2012), Higgs *et al.* (2012) Schuwirth and Van der Vleuten (2010), Kemmis (2009), Boud and Falchikov (2007), Epstein (2007), there has been no research into what form this takes within CM. The basis for my interest in assessment of professional practice is the proposition that assessment drives learning and teaching (van der Vleuten *et al.*, 2010) for better or worse, and impacts on the behaviour of both the students and the teachers. An investigation of the range and frameworks of assessment of professional practices (Pangaro and ten Cate, 2013) being applied in HE institutes of CM will help me to understand their pedagogical approaches and open up a dialogue on the nature of professional practice in CM.

   I have already looked at accreditation documents in your country and now I wish to look at:
   1.Course and module handbooks, assessment documents (structure, criteria for marking, value, etc);
   2. Attend your clinic as an observer in order to understand the context of your assessments;
   3. Interview course/programme leaders and a sample of teachers who carry out assessments.

2. **What type of participant is needed?**
   Course or Programme Leaders of Chinese Medicine professional entry courses in Universities in Australia, the UK and the USA who have been involved in developing the curriculum and thus have insight into the assessment practices within the courses.

3. **What does the study involve and what will participants be asked to do?**
   I will be asking you to:
   a) give me access to relevant documents that outline and describe assessment practice of Chinese medicine. I would like to spend time at your institution looking at the documents either hard copy or on-line.

   b) take part in a face-to-face or phone/Skype interview of 45-60 minutes at a time and place convenient to you that will be audio recorded.

    c) Agreement and arrangement for me to attend the clinic for observation purposes (I will not be recording or filming) for 1-2 weeks.

    d) recommend (and permission to contact) teachers involved in assessments of professional practice. Teachers will be sent an information sheet and consent form and invited to participate in a face-to-face or phone/Skype interview of 45-60 minutes with me at a time and place convenient to them. The interview will be audio recorded. The focus of the interview will be their experience of assessment practice. A quiet place to interview would be desirable.

4. Is participation voluntary?

Your participation is completely voluntary. You will only be included if you give your signed informed consent. Even if you do participate you can withdraw at any point and after data collection can withdraw your data without reason up to three months after interview. Please note that your identity will remain confidential and neither your colleagues nor anyone else will be informed of your participation.

You written consent will be stored in a locked filing cabinet at my University, Department of Life Sciences. Transcribed data and field notes will be kept for up to five years following publication of results.

6. How will your privacy be protected?

In order to maintain confidentiality all names of individuals, universities and places will be replaced with pseudonyms in interview transcripts and any publications arising from the research. The exception to this is reference to accreditation bodies whose information is already in the public domain. All information that might identify you such as audio recordings will be stored securely on a password protected laptop and USB and only accessed by myself. Interviews will be given a unique code retained on an encrypted file only accessible to myself. Any hard copies of documents will be kept in a locked case while travelling and a locked filing cabinet at my University and only accessible to myself. I will be transcribing the interviews and will delete all audio recordings once transcription has been completed. Participants can ask to review the transcript of the interview and ask to edit or erase their contribution up until three months after data collection.

7. What use will be made of information collected?

Findings will be disseminated through publications in scholarly books or journals, presentations at seminars or conferences. As with any research I hope it will be used to develop the CM curriculum in my own course and professional body as I hope it will with yourself and others.

8. What happens to this information after the study period?

After the study period any information linking you to the study will be destroyed. The transcribed anonymized materials and fieldnotes will be maintained for up to five years in case further analysis will be needed at any point.

9. Are there any advantages or disadvantages to participating in the study?

The final report and any publications arising from the research will be sent to participants. It is hoped that any evidence of trends and conclusions reached will help

to improve our understanding of assessment practices, the pedagogy behind them and thus help development of curricula.

I do not foresee any disadvantages and would hope that your participation in the study will be beneficial to your ongoing professional development both as a course leader and for the teachers. Even participating in the interview can trigger new ideas in relation to your own knowledge and understanding and I aim to carry out the interview with the approach of 'travelling with you', being curious and sensitive rather than 'mining' for information.

10. **Can I tell other people about the study?**
    You are welcome to discuss your contribution with anyone you choose.

11. **What if I require further information?**
    You can contact Felicity Moir at any time at the address below.

12. **Will I be debriefed at the conclusion of the study about the results or pending publications?**
    The final report and publications arising from the research will be made available to all participants.

13. **What if I have any complaints or concerns?**
    You can contact xxx who is my supervisor at the Institute of Education via email if you have any concerns about the research process or any publications.

14. **Who has reviewed this study***
    This study has been reviewed by the Institute of Education Ethics Committee and my supervisor xxx at the Institute of Education.

If you would like to participate with this research please contact me by email. I thank you in advance for your help

Kind regards

Felicity Moir
Principal lecturer and Course Leader
Herbal and East Asian Medicine Division
Department of Life Sciences
University of Westminster

Participant Consent Form

Please put your initials to each statement in the boxes

1. The procedures required for the project and the time involved have been explained to me and any questions I have about the project have been answered to my satisfaction.

2. I have read the Participant Information Statement and have ben given the opportunity to discuss the information and my involvement in the project with the researcher

3. I understand that I can withdraw from the study at any time without affecting my relationship with the researcher.

4. I understand that my involvement is strictly confidential and no information about me will be used in any way that reveals my identity other than to the researcher.

5. I understand that sections of my interview will be transcribed and used to disseminate the results of the study in seminars, conferences and peer reviewed academic or education publications, however that my anonymity will be preserved at all times.

6. I understand that my participation will be audio taped and I am aware of and consent to your use of these recordings to explore salient themes and features.

7. I understand that I can stop the audio recording of the interview at any time, and that any aspects that I choose of the recordings will be erased and the information provided will not be included in the study.

8. I understand that audiotapes will be treated as strictly confidential and will be kept in accordance with Institute of Education Research Ethics guidelines. Any raw data on which the results of the project depend will be retained in secure storage in accordance with the Data protection Act (UK, 1998)

9. I know that my participation should not lead to any potential harm or discomfort and I consent to the processing of the interview for the purposes of the study.

Participant Consent

……………………………………… ……………………………………………… ………………...

Name of Participant   *Signature*   Date

Investigator's Statement

I,……………………………………….., confirm that I have explained the purpose of the study to the participant and described any foreseeable benefits or risks deriving from his/her participation.

…………………………………...   …………………………...

   Signature   Date

Appendix 2 Teacher Information Statement

An Exploration of the Assessment of Professional Practice in Chinese Medicine Institutions in the West

---

You are invited to participate in a research project being conducted by Felicity Moir as part of her Education Doctorate at the Institute of Education, London, UK. Felicity is a Principal lecturer and Course Leader of a BSc (Hons) Chinese Medicine: Acupuncture and MSc Chinese Herbal Medicine professional entry course at the University of Westminster, London, UK.

5. **What is the study about?**
    This study hopes to explore the way that Chinese medicine (CM) Higher Education institutions in the West assess the professional practice of their students. While the assessment of professional practice in the biomedical health professions has been much debated and researched, for example, Billett (2012), Higgs *et al.* (2012) Schuwirth and Van der Vleuten (2010), Kemmis (2009), Boud and Falchikov (2007), Epstein (2007), there has been no research into what form this takes within CM. The basis for my interest in assessment of professional practice is the proposition that assessment drives learning and teaching (van der Vleuten *et al.*, 2010) for better or worse, and impacts on the behaviour of both the students and the teachers. An investigation of the range and frameworks of assessment of professional practices (Pangaro and ten Cate, 2013) being applied in HE institutes of CM will help me to understand their pedagogical approaches and open up a dialogue on the nature of professional practice in CM.

    I have already looked at accreditation documents in your country and your Course and module handbooks and assessment documents (structure, criteria for marking, organisation etc).

6. **What type of participant is needed?**
    Teachers of Chinese Medicine professional entry courses in Universities in Australia, the UK and the USA who are involved in and have insight into the assessment practices within their courses.

7. **How much of my time will be involved?**
    The interview will be 45 – 60 minutes face-to-face or by phone/Skype at a time and place convenient to you.

8. **What does the study involve and what will participants be asked to do?**
    The focus of the interview will be your experience of assessment practice at your institution:

9. **Is participation voluntary?**

Your participation is completely voluntary. You will only be included if you give your signed informed consent. Even if you do participate you can withdraw at any point and after data collection can withdraw your data without reason up to three months after interview.

Your written consent will be stored in a locked filing cabinet at my University, Department of Life Sciences. Transcribed data and field notes will be kept for up to five years following publication of results.

6. **How will your privacy be protected?**
   In order to maintain confidentiality all names of individuals, universities and places will be replaced with pseudonyms in interview transcripts and any publications arising from the research. The exception to this is reference to accreditation bodies whose information is already in the public domain. All information that might identify you such as audio recordings will be stored securely on a password protected laptop and USB and only accessed by myself. Interviews will be given a unique code retained on an encrypted file only accessible to myself. I will be transcribing the interviews and will delete all audio recordings once transcription has been completed. You can ask to review the transcript of the interview and ask to edit or erase your contribution up until three months after data collection.

7. **What use will be made of information collected?**
   Findings will be disseminated through publications in scholarly books or journals, presentations at seminars or conferences. As with any research I hope it will be used to develop the CM curriculum in my own course and professional body as I hope it will with yourself and others.

15. **What happens to this information after the study period?**
    After the study period any information linking you to the study will be destroyed. The transcribed anonymised materials and fieldnotes will be maintained for up to five years in case further analysis will be needed at any point.

16. **Are there any advantages or disadvantages to participating in the study?**
    The final report and any publications arising from the research will be sent to participants. It is hoped that any evidence of trends and conclusions reached will help to improve our understanding of assessment practices, the pedagogy behind them and thus help development or curriculum.

    I do not foresee any disadvantages and would hope that your participation in the study will be beneficial to your ongoing professional development. Even participating in the interview can trigger new ideas in relation to your own knowledge and understanding and I aim to carry out the interview with the approach of 'travelling with you', being curious and sensitive rather than 'mining' for information.

17. Can I tell other people about the study?
You are welcome to discuss your contribution with anyone you choose.

18. What if I require further information?
You can contact Felicity Moir at any time at the address below.

19. Will I be debriefed at the conclusion of the study about the results or pending publications?
The final report and publications arising from the research will be made available to all participants.

20. What if I have any complaints or concerns?
You can contact xxx my supervisor at the Institute of Education via email if you have any concerns about the research process or any publications.

21. Who has reviewed this study?
This study has been reviewed by the Institute of Education Ethics Committee and my supervisor xxx at the Institute of Education.

If you would like to participate with this research please contact me by email. I thank you in advance for your help

Kind regards

Felicity Moir
Principal lecturer and Course Leader
Herbal and East Asian Medicine Division
Department of Life Sciences
University of Westminster

Teacher Consent Form

An Exploration of the Assessment of Professional Practice in Chinese Medicine Institutions in the West

Please put your initials to each statement in the boxes

1. The procedures required for the project and the time involved have been explained to me and any questions I have about the project have been answered to my satisfaction.

2. I understand that I can withdraw from the study at any time without reason up until 3 months after the date of the interview __/ _ / ___

4. I understand that my involvement is strictly confidential and no information about me will be used in any way that reveals my or my institution's identity other than to the researcher.

5. I understand that my participation will be audio taped and I am aware of and consent to your use of these recordings to explore salient themes and features.

6. I understand that I can stop the audio recording of the interview at any time, and that any aspects that I choose of the recordings will be erased and the information provided will not be included in the study.

7. I understand that sections of my interview will be transcribed and used to disseminate the results of the study in seminars, conferences and peer reviewed academic or education publications, however that my anonymity will be preserved at all times.

8. I understand that the audiotapes will be treated as strictly confidential and will be kept in accordance with Institute of Education Research Ethics guidelines.. Any raw data on which the results of the project depend will be retained in secure storage in accordance with the Data protection Act (UK, 1998)

9. I know that my participation should not lead to any potential harm or discomfort and I consent to the processing of the interview for the purposes of the study.

Participant Consent		
Name of Participant	*Signature*	*Date*

Investigator's Statement
I,………………………………………….., confirm that I have explained the purpose of the study to the participant and described any foreseeable benefits or risks deriving from his/her participation.

……………………………………..   ……………………………..
      Signature   Date

Appendix 3 Interview Questions to Course Managers

Research Questions	Interview Questions
General context of the individual manager	Tell me about yourself and what role have you had in the development of the programme and specifically the assessments.
What do the managers understand by the stake holders and their impact on assessment choices	What and who were the key influences on the design of the assessments? (e.g. accreditation board, higher education in your country, experience of the educators in the institution?)
What is the relationship between the management and the supervisors	What is the structure of the teaching group? How are teachers employed? what contracts do you/they have? What teacher training programme do you have? expect? How do you train clinical supervisors and clinical assessors – are they the same? How are clinical teachers selected? are they the same as classroom teachers? What are the challenges/ issues that seem to come up for those assessing practice? How do supervisors see their role as teacher and assessor; as formative assessor and summative – what are the differences.
How do the managers understand the issues behind the application of assessment instruments to practice-based assessments	Tell me about the thinking behind assessments generally and of clinical practice. You have learning outcomes in each unit of study? Do you have criteria for each assessment in each unit? Who has developed those? If so how closely do unit learning outcomes / assessment outcomes align with course learning outcomes need to align with course LO's? Where has this requirement come from? Tell me about the final exams – structure, marking criteria, etc. How are they constructed? What sorts of assessments do you use? What have you tried and rejected? How do you moderate clinical assessments? How closely do students results in clinical practice assessments align with results in other areas of knowledge and skills – do you see a difference? How much do you think personal judgement comes into assessments?

Using the idea of 'failing' how are the criteria administered	What criteria are required to be passed or if a student fails any one criterion would they fail overall??
	How do you deal with different levels of student ability in clinic;
	
	       how do you deal with 'fitness to practice'
	       how do you deal with the marginal performer;
	       if a student fails what happens next
	Do you have external examiners for clinical assessments?

Interview Questions to Supervisors

Research Questions	Interview Questions
General context of the supervisor	Tell me about where you trained? How long have you been at xxx? How long as a clinical supervisor?
How do they understand their development of their role	Tell me about how you came to be a clinical supervisor? Where would you say your style of supervising comes from? What/who have been the key influences? What do you think are the key skills of a supervisor?
How do they understand assessment	Tell me about the decision making process in clinic in terms of deciding a diagnosis and treatment between you and the student? How do you go about giving feedback? What are the issues of students taking on feedback and learning from it? How do you deal with weaker students? Any special techniques? How much do you find students able to critique their own work? It is said in education that "students learn from trial and error and need to feel comfortable enough to have a go" – how do you understand that statement in relation to clinical teaching?
What do the supervisors understand about learning outcomes	Tell me about how you go about assessing students. What are the key things you expect from students as practitioners? What is the most difficult part of assessing students in clinic? Have you ever had to fail a student? Tell me about it? It is sometimes said "I do not know how to define quality but I know it when I see it". How do you understand that statement?
What is their experience of using the assessment instrument	Tell me about using the marking criteria grid. Do you come to a grade or a pass/fail? What sorts of assessments are you involved with other than clinic? Do you find either easier or harder?
What community of practice do they identify with	As a team what sorts of discussions do you have around assessment? If you had a problem in clinic who would you talk to about it? Who do you talk to generally about clinical teaching?

Appendix 4 Analysis of Accreditation Documents

First Analysis Codes	Second Analysis Codes
Name of accreditation board	**Philosophy**
Authority/ Statutory regulation	**Community**
Membership of accrediting body	**Communication**
Documents	Public
Level of award	Accountability
Length of programme	Management
Standards	Resources
Core curriculum / Content	Evaluation
Clinical Hours	Mapping
Nature of clinic/resources	Hours – inputs, outputs
Competencies/learning outcomes	Supervisors training
Theory and practice	Teacher qualifications
Knowledge/skills/aptitudes	Variety of supervisors
Assessment and Fairness	
Measurement	
Themes	
alignment with regulators	communication
inputs and outputs	teachers and teaching
atomisation and measurement	resources
assessors and fairness	the (lost) discourse of CM

Example First Analysis

	ACAOM	CMBA	BAAB
Competencies / learning outcomes	7 professional competencies follows what is expected through a CM treatment and broken down into further multiple competencies: collecting data, formulating a diagnosis, determining a treatment strategy, performing treatment, assessing effectiveness, complying with codes of ethics and practice, aware of red flags & integrating biomed. (2019 document is substantially different)	Capabilities broken down into 6 standards (professional & ethical conduct, communication & collaboration, reflective practice & professional learning, quality & risk management, CM practice) and further defined as learning outcomes (integrated)	6 themes broken down into standards: context, diagnosis & treatment, communication & interaction, safety, professional development, business management - and further into multiple integrated competencies written as learning outcomes

Theory and practice to be integrated	*didactic, practical and clinical appropriately integrated*	integrating theoretical knowledge and clinical practice of Chinese medicine throughout the program *designing an integrated, structured clinical education and placement program that provides each student with experiences (including simulated learning and opportunities for inter-professional learning) across the scope of practice expected of entry level Chinese medicine practitioners*	*Integration of theory and practice must be central to the curriculum clinical practice must be central to and integral with the entire programme*
Measurement	*Criterion 7.02 … A variety of assessment measures must be systematically and sequentially applied throughout the professional program in acupuncture/ Oriental medicine. Assessment processes must measure student performance in the professional competency areas in accord with the outcome expectations as outlined in STANDARD 8 and the achievement of program educational objectives*	*The assessment team will want to see evidence that the Chinese medicine program learning outcomes and assessment cover each standard and that students demonstrate all of the standards during the program.* *To meet each standard in Field 6, the education provider must provide evidence that shows where each of the standards are covered in the program and how each of the standards are assessed.* *To do this you will need to identify the learning outcomes of the program and map those learning outcomes to the standards*	*SETAP 6.4: Assessment methods must be employed that measure the learning outcomes and skills required to practise safely and efficiently.* *SETAP 4.1: …you will be asked to show how your learning outcomes meet the BAcC's Guidelines for Education by cross-referencing your learning outcomes to that document*

Example Second Analysis

Phil-osophy	1 use	1 use	13 uses
	The statement of purpose and goals should reflect the purpose for which the program was founded, the <u>philosophies</u> it represents, the community in which it is located, the constituencies it serves, the needs -- social, cultural and material -- of its community and clientele, and the institution's resources -- human, physical, and financial. P5	*What is your <u>philosophical</u> stance towards the development of clinical skills and how is this reflected in the sort of clinical studies' workbooks/portfolios you have created to give structure to your clinical education program? p21*	educational <u>philosophy</u> and the relationship of this to healthcare in the country and the local community p23 includes the Board <u>philosophy</u> towards the process of accreditation **practice-led**: *rooted in the <u>artistry</u> and skills of acupuncture as an empirical and practical professional activity informed by theory and <u>creative of theory</u>, and recognizing that, as acupuncture is a practice-based profession, acupuncture teachers will normally remain engaged in practice p7*
Comm-unity	8 uses	5 uses	4 uses
	A. The program shall encourage the academic discourse between faculty and students which results in the development of an academic <u>community</u> that will enrich and advance the profession, and contribute to the development of future generations of practitioners, faculty, researchers, clinical supervisors, and leaders of the profession. Others in relation to patients	*How do you engage with the broader Chinese medicine professional <u>community</u> and the health sector?* *Do you think the program is well-regarded by the broader Chinese medicine professional <u>community</u> and health sector? If so, what makes you think it is? If not, what could be done to address this issue? P9* Others in relation to patients	*It is essential for students to know and to understand the importance of working with and as part of the wider health and social care <u>community</u>, p32* *Collaborative: creating opportunities for shared learning with other healthcare professions, recognizing a common purpose in the wider <u>community</u> p7*

| Commun-ication | '*Provision must be made for regular and open communication among members of the faculty and between the faculty and administrative officers of the institution*' p26 | No use of the word communication other than what students will learn to do in relation to patients

Want staff evaluations but no recommendations on quality of communication | '*The Board will also investigate how communication is maintained with clinical supervisors and seek evidence of the way that these members of staff are appointed and appraised, as well as how they are supported and encouraged to take part in developing and changing the programme in order to ensure that clinical learning outcomes, levels and progression are consistent with the rest of the programme*' p35 |

Appendix 5 Analysis of Institutional Documents

Authenticity
Own in-house outpatient clinics (some students also did placements in hospitals and in China)
Busy clinics with waiting lists
All off- the- street for easy access for public
Receptionists and clinic staff and supervisors and students worked together in relation to patient care
Students responsible for patients with close supervision (more removed as progress)
Supervisors went into treatment rooms with students giving immediate feedback
Similar set ups with discussion rooms and either single treatment rooms or beds divided by curtains
Discussion room - with groups of students or sometimes singly with a student
Similar paraphernalia of beds, white coats, needles and moxa, charts on walls

Learning outcomes and measurement
All clinic modules had multiple assessments including case histories or reflective journals as well as the WBA
Assessment instruments contained learning outcomes written in behavioural language
Some same as accreditors and some different
If drill down learning outcomes similar
Multiple outcomes anything from 20-25-97
Numerical mark expected at end but instruments differed re space for a numerical mark
Some ask for number, some competent/not competent, or fail – excellent (5 grades)
Instruments for formative and summative assessment at different times of student year
All had attendance logs and stipulated hours to complete; some had competency logs to be filled in daily
Each institute different values for the WBA – some very little and others more – some led to full marks; only one had differentiating criteria
Language devoid of any philosophy of CM

Language in documents
Mission statements contain language of CM;
Descriptions of role of clinic in course handbooks (language relates to CM philosophy, and discourse of praxis and relationships)
Clinic documents talk praxis and values and community; UK doc seems more to emphasise CM
Link the role of clinic to professional learning

Comparison of descriptions of clinic in handbooks – example

Module descriptor or clinic handbook	describes the expectations of students in the clinic emphasising the role of the student in their social relationship to the clinic and the	is extensive with overall subject objectives that cover didactic classes as well as the clinic placement. There is a description of the	is quite discursive and presents the clinic as a place where practice is happening and so emphasises authenticity. It describes the role of the supervisor and their

	supervisor. The role of the supervisor is also described with the language emphasis being on the legal relationship of the student to the supervisor and that the final responsibility for patient care rests with the supervisor. The student must gain signed approval for all treatment plans.	clinic as an authentic setting that will mimic the graduates practice, and describes the relationship of the student with the supervisor. The language could apply to any health care setting and is not using any specific language that could be identified as CM.	relationship with the students and describes the clinic environment in the language of CM.

Comparisons of Learning Outcomes across institutions - examples

N.B. It is difficult to make comparisons in relation to general headings but when broken down to sub-competencies most of them can be mapped across all courses.

Competencies	USA – written as LO's following different structure from accreditation documents (formative (but a mark given) written differently from summative	Australia – written as competency descriptors & broken down into performance criteria Works through the linear aspects of a consultation	UK – written as performance areas through the linear aspects of a consultation. The clinic log lists other areas and is written as LO's
Communication Skills	*Counsels and educates patients and families Demonstrates ability to effectively collect, evaluate, present and explain clinical data, OM diagnoses, treatment principles, and treatment plans to supervisors and peers. Demonstrates the ability to communicate effectively and collaborate with patients, supervisors, staff, and other healthcare professionals Models good communication skills for observers,*	*Patient Communication and Rapport: to ensure a professional attitude, education of the patient as to the benefits of Chinese medicine. Includes directives to the patient regarding the frequency of treatment Communicates with Clinical Supervisor: to ensure safe and appropriate treatment is administered Liaison with Fellow Student Practitioners: so as to promote peer responsibility and the smooth running of the clinic*	*Contact with patients Looking and listening skills Developing rapport, sensitivity to patient issues*

	assistants and associate interns. Exhibits skills in handling challenging patients and interactions			

Description of assessment instrument – example

The ethos of the assessment form was a psychometric grading of performance as determined by the supervisor against a set of questions or criteria that were written as outcomes. Knowledge, skills and aptitudes were to be demonstrated.
This document did not specify but it looked as if all competencies must be passed. In this assessment there was a differentiation of criteria into subjective and objective with the subjective (still graded) including the manner by which the student managed the consultation; one assumes the subjective implies an awareness that this is a judgement being made by the supervisor. What does this imply, however, for the other criteria being objective?

Students completed case notes (which I am interpreting as product) were also graded against a set of outcomes.

Appendix 6a Cross Case Management Interviews

Issues			
History and Organisation	Development and evolution of college from a small organisation founded by practitioners, to a bigger corporate owned entity.	Developed as acupuncture only programme as a private school and moved into University.	Started as a private college by a group of practitioners as a PT course that has moved to fulltime. It is registered as a charity.
Accreditation and Validation	ACAOM is professional programme accreditor and also has a different institutional accreditor	Course already set up before Statutory Regulation so not modelled around Accreditation document	The degree is validated with a University and course and institution accredited by the BAAB.
Communication and Decision Making	Has a curriculum committee within the faculty governance and a staff and faculty advisory committee (academic directors, department directors, FT faculty, adjunct faculty, sometimes students). Meeting with supervisors once per trimester.	Meetings are held a few times per year formally but a lot of informal meetings between clinic supervisors and course directors and clinic managers who also work in the clinic.	Board of Studies will sign off changes to modules and to assessment briefs. The clinic team meet up a few times a year to discuss student final grades and also review assessment processes.
Criteria and Mapping	Mapping of LO's is an important area and something being done by faculty and course leaders. Programme LO's map onto Course LO and Course LO's map onto assessments and have % requirements	University currently looking at consistency in Course and Subject Learning Outcomes. Assessments map onto learning outcomes. Assessments had marking criteria grids but not clinical assessments.	Content maps to Learning outcomes and maps to assessments within modules and also from course learning outcomes into module learning outcomes. Clinical assessments had well developed marking criteria grids.
Employment	Clinical supervisors do not have contracts - they are employed on a sessional basis per trimester	No new teachers for a long time. All clinical teachers were tenured.	Clinical supervisors are on contracts and teach one day per week on average in clinic - part of the contract is to attend meetings.
Teacher Training	There is no teacher training programme for new staff and no on-going CPD for teaching. They prefer new teachers to come	No formal teacher training programme.	No teacher training programme for clinical supervisors; have had various one day events in the past

	in with teaching experience but that is rare		
Student Choice with Clinic Supervisors and Evaluations	Students can decide whom they want as supervisors. Students evaluate the courses and the clinical shifts.	All students will be seen and assessed by all supervisors. Modules are evaluated not teachers.	Students chose when they come to clinic; modules are evaluated not teachers
Assessment	In clinic there is a (summative) mid-term feedback against an assessment instrument and supervisors are allocated time to do this. There is an on-line written assessment at the end of each shift against an instrument – not allocated time for this.	Two forms in operation. New form was written to meet a different target group and as part of National registration.	5 marking categories from Excellent - Needs Development came from discussion across all courses. Some decisions for assessments come from external examiners.
Structure of the Clinics	The clinic was on the ground floor directly open to the public and included a reception area and waiting room, a number of individual clinic rooms, a herbal medicine dispensary and discussion rooms. Separate from academic offices and classrooms	The clinic was on the ground floor directly open to the public and included a reception area and waiting room, a number of individual clinic rooms, a herbal medicine dispensary and discussion rooms. Separate from academic offices and classrooms	The clinic along with all other facilities was housed in one building that opened onto the street. Patients and students and staff entered through the same door. Patients would walk past classrooms to get to the clinic area. Reception area, discussion room, a number of treatment rooms, herbal medicine dispensary.

Appendix 6b Codes and Interview with Manager

<u>CODES</u>

Clinic as resource – costs

Psychometrics and learning outcomes – some seem pro, others aware of issues

Grades or pass/fail – grades expected

Mapping across LO's and assessments – standardized process

Criteria – only in one institution

Grade inflation and failing-to-fail – all (most) students get an A.

Didactic classes determine failure – issue of objectivity?

Objective – subjective to do with evidence

Nature of competence – given more time in order to meet requirements

Know the students - justification?

Inputs and outputs – both operating

Stake-holders and assessment tools - institutional choices

Numbers of supervisors - varied

Recruitment/Training/induction of supervisors – where is education?

Contracts – tenure or short contracts

Communication with supervisors – meetings, but how many and how long

Community – what time given to this?

Time – with students, for meetings, to carry out assessments, casual encounters

Accountability – supervisors not aware of final marks in 2 institutions

Litigation and subjectivity

Observations

In-house

Location of clinics

Busyness of clinics

Places of work

Staff/student ratio (link to accreditation requirement)

Hands-on - supervisors work closely with students and patients

Codes – examples

Codes	Quote	Interpretation
Failing students And	Tell me about how many students pass/fail clinic? *Oh pretty well everyone does [passes]*	Input-output – all institutions talked about number of hours as significant in relation to learning.

Inputs	*because we have such extensive training. 325-326.* *…. they fail academic before they fail clinic. Clinic is very difficult* [to fail] *357.*	Contradicted by another manager who considers hours not indicative of competence. Another institution says similarly Knowledge informs practice but not necessarily seen other way around.
Stake-holders and assessment instrument	Because I was programme director I had to demonstrate evidence that our students met these outcomes and because our accreditation depended on it so I jumped in with this document.461-463	Standardisation comes from accreditors as stakeholders who hold the power – not supervisors. Contradicted by another manager who thought statutory changes had not impacted on course. Another institution also pursuing standardised LO's but not as an accreditation requirement.
Numbers of supervisors	So we've a great strength in the diversity that we've got, with our five different… six days of the week, five supervisors, that's really great strength, so the students get to see a lot of different skills, but within that diversity, that's also, the need to get everyone to conform to certain basic… to achieve that they have meetings, but the problem we've had in the past, it's the old thing, it's such a busy environment, that those meetings tend to be a lot of nuts and bolts. They have to discuss every student. 375-381	Variety of supervisors as a strength in tension with training and equality on interpretation of learning outcomes. In another institution no's of supervisors relates to employment The busyness of clinic environment limits communication. This institution does discuss each students result as a team – good communication leading to commonality of interpretations.

Interview with Manager – example of coding

It's in the clinic manual. It is a loose grading criteria compared to the classroom, is like part of it is based on attendance, no brainer there, and part is based on paper work they do, activity reports, that is somewhere about 15% or 20% depending on level of students and a lot is kind of subjective component that the supervisor grades on , more often than not the supervisors tend to grade very liberally in that regard which is OK as the supervisors are in there a lot and they know if a student is not doing , they are in there with the student watching their needling and listening to their presentations, so I mean, I would even say it should be pass-fail because essentially they are graded as a pass-fail because if it is either a subjective component that the students get an A on unless there is something really amiss, so just that part, so the only reason you would give an A minus in the clinic is because of attendance issues which is objective or because they have not handed in the appropriate paper work which is objective. So the objective part is where they will lose grades but they will still pass. You know a C is a pass

so what you are saying is for a student to fail they would have to be not present or not handing in the paper work

What about therapeutic relationship, patient relationship, would a student fail? Lack of rapport with patients?

Where patients don't come back? That's very unusual first of all. That is rarity. It makes sense it is rare. The reason why it is rare is that the students spend a lot of time on the programme, it is a lot of money, a lot of effort, so you get students who for the most part are very serious and for the most part they have been vetted out before they become a clinic intern. Somewhere in the observation or assistant level. What vets them out is they can't pass the comprehensive or the comprehensive exam. And it is not the clinic work, which is very subjective, but the didactic class.

Appendix 7a Supervisors Individual Stories

Supervisor 1 has been supervising for many years. She is a Chinese national and talks about the difference in approach from her experience in China that she needs to take with students, a more step-by-step approach. This seems to imply students come to clinic with less knowledge of how to elicit information and put a diagnosis together. Similarly to Chinese students, students here she says <u>want to learn</u>, they want to ask questions but some are high achieving and some are poor in their studies. She works out their level by the questions they ask and then uses language at that level.

She thinks you need to have <u>knowledge</u> and a high professional level and she is the only supervisor who talked about including a Western diagnosis as well as Chinese. Primarily however her emphasis was on <u>loving</u> students and teaching and being happy 'pulling out your knowledge and <u>knowing your students'</u>. She says the students work hard and they 'really feel <u>I love teaching them'</u>. She describes <u>working as a team</u> and describes <u>knowing the students very well</u> because she spends <u>so much time</u> with them.

She <u>goes into the treatment room</u> with the students to ask more deep and broad questions of the patient and with the student makes the diagnosis. It is important to see the patient as well, as you do not want the student repeating mistakes.

While she is aware of the <u>assessment forms</u> and when they should be done, she talks about <u>assessing the student every time</u> that translates as <u>regular feedback.</u> If they are poor in an aspect she gives them work to do and checks the following session especially diseases that are not common and so may not have been covered in class. Because she gets to know the students so well she does '<u>not want to fail any student'</u> and would <u>feel 'embarrassed'</u> to do so.

> *We build up a personal relationship, we love the student and student love us.*
> *225-226*

Only students with very bad behaviour would she be giving a low mark to and could <u>only describe one student she ever failed</u>, who did not 'have time to study' because of family commitments. She had an interesting view on what clinical training is about

> *for education as we don't want to give the student like penalty but education.*
> *So let the student realise he is not doing good and he need to do better 236-*
> *238.*

She works with students until they improve. Even students who do something dangerous such as leaving a patient with needles in, she sees as learning. The student never did it again. In all her time working in the clinic they have never had an accident.

Supervisor 2 always wanted to be a teacher to 'share' what he had to give. He sees the students as '<u>very happy, willing and responsible in the clinic'</u>, 'relaxed and enthusiastic', and so does not have to play 'police officer'. He tries to see the students perspective of learning to be in practice and a key role is to teach the

students how to 'interact closely with people', to teach the 'patient practitioner relationship' as 'from that spills courtesy, respect, sympathy, empathy'. He expects students to listen to him as he has something to share and likes to create a pleasant atmosphere. What he wants to 'cultivate' is the student's '<u>intuitive ability'</u>, to 'draw it out' of them. He wants them to develop their relationship with patients:

> *to have <u>insight</u> into that beautiful wonderful human being who is in front of them who they are trying to treat 274-275*

<u>Students sign up for his shift</u> he suggests, because he has a different perspective from others, he is 'going to get a good diagnosis and interesting treatment for them to do'. They are not in a hurry and have time to reason and discuss the diagnosis together. He wants the experience to be positive for the student and will work with the student at the level he finds them trying to give equal attention to all. He wants to 'build them up' not 'chop them down', something he thinks other supervisors do. He considers however that this approach may be constructive and maybe both approaches are needed. He feels he does not really get to know the students until a good way into the shift as they take time to find their style and so the mid session feedback is too soon. Some students he feels have 'conceptual limitations'.

His key expectations of the students are to <u>turn up on time, to do the paperwork, follow the rules, to be 'courteous to their patients and treat them as the most important thing'.</u> He feels he is <u>'lucky'</u> as he has never had any difficult students. He <u>has never failed anyone clinic</u> as 'they like clinic they want to be part of that experience.' He does not say negative things. He is aware of grade inflation but says:

> *I can't see why other than personal disagreement with someone why I would not <u>give them a solid good grade</u> 361-362*

Where students get to at the end of the course is only a beginning he explains. Some students are slower and need more time but it is going to be 'rocky' out there. They do not have the same amount of training as MD's so no need to be hard on them.

Appendix 7b Codes and quotes Supervisor Interviews

Codes	Key words and phrases
Training	Graduate of own course x 7 Graduate from China x 5 Years since graduating 5-38 Years as supervisor 2-22
How learned to supervise	From own supervisors x9 Previous training x5 Shadowing x5
Teaching methods	help students make connections with theory, set stages for students, breaking it down, helping students to focus on key issues, wanting them to think and be independent, sensitive of the students learning style, giving them free rein with control, guide, empowering students to make best choices while making sure patients getting best treatment, a balancing game, aware of need to support dynamic with patient
Supervisors view of self:	ability to tell a story, being supportive, to facilitate, advising, giving confidence, loving students, being patient, to 'cheer lead' and keep them [students] on the right track, trusting students, not condescending or pushy, being honest, not criticise or constantly correct, cultivate insight, belief in intelligence of students, being sensitive to the students situation (including financial), aware of cultural differences; run multiple rooms; personal perspective; passion; liked by students; style of supervision was different in some cases or similar from other supervisors; feedback to students important; don't follow rules; conduit
Relationship with students	love of teaching, patients, medicine; love students and love teaching and happy to pull out your knowledge and knowing for the student; knowing the students x2 (debt, family life, finances) the students really feel I love teaching them. I have this feeling - so we work as a team; a collaborative relationship when in a clinical environment, rapport
Issues of supervision	knowledge and skills & experience, superlative diagnostic skills, having a practitioners eye; good grasp of the basics; confidence, lack of anxiety; time management; to be good at explaining; to give information; to give guidance; to have ability to explain more than knowledge; to question & discuss students thinking; to show equality; to show to have leadership; patience; to check patients to verify students; to check techniques; to expect/not expect research; to work with time pressure; to work with student & patient; to know level of student as teaching changes with level of student; to work with student as a team; to be fair
Supervisor role to protect patient	The supervisor is the leader and has to go with the student and the patient; you cannot let them make a mistake again and again. If mistakes are made then that is why the supervisors are there and their peers are there to double check against mistakes. A supervisor plays a shadow, a protective shadow in there; not overpowering, not take their motivation but be there to protect the patient A collaborative relationship when in a clinical environment I will support them in any way (as long as) patients are safe.
View of	wants critical thinking; training students for future work/practice;

students	positive view of students; supportive; they want to learn; knows the students; understands students situation – feelings, personal and financial, try harder; all want something different (and not always what is in assessment instrument): knowledge, good interaction with patient, clean needle, reasonable diagnosis; proactive, timekeeping, to best ability, boundaries, insight; negative aspects (rare) – arrogance, challenging re personalities; students responsible and get things done; students happy, relaxed, enthusiastic; students open and receptive; students open as here to learn; I give students full latitude
Unique views	Honesty and dedication They need to have a loving and caring heart; (sensitivity, empathy, sympathy) Looking for intuition and insight of their patients as wonderful human beings Once they get the click they get enlightenment
Community:	Differences across institutions: practitioner, educator, academic – all see themselves primarily as practitioners; meetings rare in one and more in others; communication with other supervisors in 2; too little time; attitude to other supervisors positive; talk to practitioners outside institute, close community talk to each other, or regular meetings & discussions with all teachers and managers
Time	Students: to organise, managerial, run the clinic shift, time management x2, to be able to treat more quickly, staying on schedule, Assessments: Does mid term evaluation with students as has the time; sits with students to do mid-term evaluation re strengths and weaknesses; need enough time to give proper assessment; problems with using forms is not enough time; long form not working - too much hassle, takes too much time and too busy in clinic With others: no time to talk; not much social (time) with other teachers; sometimes meet at lunchtime; not much discussion with other supervisors since moved In one institution little formal meeting time – related to comments about lack of time
Assessment – failure to fail	difficult to fail clinic; never failed anyone (only 1 example given); would fail before coming to clinic; major incident does not lead to failing; failing would need to be something ('egregious') or not doing paperwork or not turning up i.e. objective measures; cannot fail as will do again until meets competence; do not like to fail students; no external scrutiny of clinical assessments in 2 institutions; give students extra time if think will fail in 2 institution; would stop them at mid term evaluation if so bad
and grade inflation	Students expect A grade; students grade themselves A no poor grades – students like clinic On the form there are marks for different elements at mid-term but for the final you are just meant to give an A. I think they all get A's. I don't know how it is meant to work Students think an A- is a low grade. Generally mark at good or excellent;
Assessment – marking and grading	someone else compiles final marks so supervisors do not know outcomes 2 institutions; in other they discuss final marks as a team; have marking schemes but do not follow process or seen as irrelevant or thinks process

	cannot work or do not understand process; grading criteria written specifically but used loosely; students challenge grades 1 institution; school will uphold poor marks; critical of system 2 institutions; relationship with students impacts on assessment; aware of stakeholders influence on assessment process (x2); too busy to use system as intended; grade inflation in one institution and range of marks in other
Assessment instrument	Accepts that rubric must be as is because of stakeholders but does not see them as related to CM, principles of CM Does not like the numerology in new rubric; could not take criteria seriously; could not answer with veracity Would prefer grading based on meeting milestones with Y/N rather than a mark School try to make form objective but still difficult Form very hard (to use). Long form too big for a one off case but OK if over a semester. Short form supposed to be on one patient but it is of general practice as student nervous with the one case so bases marks on whole performance. Long form not working - too much hassle, takes too much time and too busy in clinic Sees the performance criteria as a whole; we follow the students a whole year, we know them well, do not go through each single thing; fills in according to assessment over year. Marking as having a subjective component Different from marking a didactic class – clinic is about actual practice, the doing What you see on paper isn't always what you get. Doesn't know about the assessment instrument but aware of a list for the assistants and observers which she checks weekly So I think the way it's done is probably the fairest. Does not use rubric in as too complicated, too many words Formal rubric and clinic log don't match up well Formal rubric is not used; uses clinic log; writes detailed commentary on clinic log;
Didactic – clinic learning	hours in clinic; skills & knowledge developed before coming to clinic; differences between supervisors; keep working with students learning until pass; don't do assessment until sure to pass – feedback and more time given; clinic assessment different from written assessment (all in the mind)
Formative assessment	Feedback is done in the moment rather than on paper Gives feedback directly to student not on paper. Giving feedback all the time in the clinic The supervisor is the leader and have to go with the student and the patient Assessment is on going every day Assessment not always formal Formative assessment goes on every week, and with every patient; Positive feedback
Expectations of students	Likes to assess on what the student knows not what they don't know; rewards for strengths How do you measure devotion, attentiveness Grades on process of learning, process of building skills Will not grade on personality Interns must be present, attired correctly, respectful, performed to best of ability

	Look at them as a whole. Knowledge and how they interact with the patient Interaction with patient is hard to grade Relationship, very hard to (grade) impression energy is good you are motivated and what you do exceeds expectations and some don't get motivated and energy not good so don't perform Hardest bit is to be detached from personal, individual preference, less subjective; constantly need to be fair; How to merge reality and marking criteria because students show strengths and diversities all over the place; criteria cant cover everything; Difficult bit is personalities and interaction with patients Assessment is supposed to assess achievement not effort but gives extra marks for effort