Illuminating the Assessment Practices of Teachers in NZ Outdoor Education Tertiary Programmes

A thesis presented in partial fulfilment of the requirements for the degree of
Master of Education
(Adult Education)

At
Massey University
Manawatu
New Zealand

Stephen Raymond Chapman

2011

Abstract

The purpose of this research study was to illuminate the assessment practices of tertiary outdoor education teachers in the polytechnic sector. Outdoor education as a theoretical model and practice was quickly revealed as a contested space. One of the main issues in contention was the place of outdoor recreation, adventure and risk within outdoor education. There appeared a strong move by academics (M. Brown & Fraser, 2009; Irwin, 2010; Martin, 2005a) to reposition outdoor education as environmental and sustainability education.

The data was gathered through a questionnaire, observation of teachers as they assessed students, and from interviews with teachers. Using an illuminative evaluation approach, three themes were identified: teacher perceptions of assessment, the focus of assessment, and the practices used to make assessment decisions.

Teachers were generally highly skilled outdoor education practitioners, however, there were indications that there were gaps of understanding of theoretical assessment concepts. Teachers seemed to find summative assessment challenging but they routinely used formative assessment to promote learning and worked hard at providing quality opportunities for learning. Outdoor recreation in outdoor education was not contested by these teachers, although there was an over-emphasis on assessing technical outdoor pursuit skills, which resulted in a lack of assessment of other less tangible aspects of outdoor education. The use of assessment criteria was common practice although there were questions about how clearly these described the standard. The role of professional judgement in assessment decisions were treated with suspicion because it was seen as too subjective. However, it became clear that professional judgement was an essential aspect of their assessment practices. Criteria and professional judgement revealed themselves as mutually inclusive.

In conclusion, while teachers seem very skilled at providing quality-learning experiences, there appear to be opportunities to further develop teacher assessment practices, for example, through exploring approaches that assess less tangible aspects of practice such as leadership, judgement and decision-making.

Acknowledgments

The research process has meant many hours absorbed in books and articles, lost in thinking and spent writing on a computer. Thankfully, many friends were interested in this research and gave their time freely to discuss and critique this study. For this, I am truly grateful. Their insights, challenges, and support have helped the long and at times seemingly interminable journey for this research to become a reality. They deserve to share some of the credit for this research.

I am indebted to my supervisors, Gloria Slater and Dr Linda Leach for their support and help that they both gave throughout the research. They always gave critical feedback in a manner that was clear and supportive, beneficial for my development as an academic researcher and invaluable for this research thesis. They pushed me to do the best I could. It has been a pleasure working with them both.

This research involved travelling around New Zealand to observe and interview many specialist tertiary outdoor education teachers. Thankfully, both Massey University and Christchurch Polytechnic supported this research with funding grants to make this possible.

Finally, I would like to express my thanks to Massey University for realising that the Christchurch earthquakes were extremely disruptive, and a serious impediment to quality study and research. The offering of time extensions during stressful times allowed me to bring this thesis to completion.

Contents

Tables

The structure of this thesis

This thesis consists of seven chapters, a list of references, and a number of appendices. The topic of each chapter is briefly delineated below.

- **Chapter 1** acts as a general introduction to this research study by discussing a historical perspective on outdoor education assessment in New Zealand. This is then followed by perceived issues in the assessment of outdoor education, leading to the clarification of the research question and sub questions.
- **Chapter 2** describes the outdoor education sector and its complexity, while placing the study in a context. This chapter also provides a literature review around issues that arise from the context.
- **Chapter 3** reviews literature on assessment.
- **Chapter 4** outlines the research methodology and design of the research study, describing the research methods used and the limitations.
- **Chapter 5** reports the findings.
- **Chapter 6** discusses the findings.
- **Chapter 7** concludes the research, makes recommendations, suggests future research, and offers several concluding thoughts, including a reflection on the research methodology.

Following the final chapter, the references are presented along with the various appendices, which are numbered sequentially in the order they are referred to in the body of the thesis.

Chapter One: Introduction

An outdoor education teacher requires a wide range of diverse knowledge, skills, attitudes, and judgements (Miles & Priest, 1990) and various authors (Buell, 1983; Holt, 1974; Priest & Gass, 1997; Swiderski, 1984) have attempted to define these capabilities but with limited agreement. Historically within New Zealand outdoor education, training and assessment have occurred through some form of on the job learning. Competency based assessments were developed throughout various organisations and used in isolation from each other, creating varying standards around New Zealand schools, outdoor education centres and voluntary organisations. This happened because no national qualification or award standards existed. Outdoor teachers were simply 'outdoor enthusiasts' who usually held an under-graduate degree or a trade, had specialist outdoor pursuit knowledge and skills, and acted as specialist instructors, assisting classroom teachers (Lynch, 2006).

According to Lynch, and Lugg and Martin (as cited in Boyes & Zink, 2006) outdoor education within New Zealand and Australia is still predominantly taught by enthusiastic teachers many of whom appear to have developed their outdoor skills separately from their teacher training. If this is so, a variety of standards are accepted within schools, outdoor centres, and voluntary organisations. Lynch (2006) clearly summarises this: "Given the importance of ensuring pupil wellbeing while engaged in school activities, teacher training has not had a high priority. Overall teacher preparation for outdoor education has been inadequate and it has continued as a thorny issue" (p. 118). Furthermore, whenever there is an accident in the outdoors the competence of the teacher is likely to be questioned. Standards, qualifications, and experience are likely to be critical aspects for safety, quality experiences, and learning in an outdoor context.

Atkinson (2009) argues that standards facilitate the development of foundational competencies and help to ensure the use and development of current accepted professional practices. However, the fact teachers have outdoor competencies does not ensure they will make good decisions or teach well. Gaps can remain regardless of skill, knowledge and experience. Incidents or fatalities could still occur, as could poor teaching practices. Standards and training are, however, likely to help reduce

the occurrences of these negative experiences. From my reading, teaching experiences and discussions with peers I initially identified a number of key issues for my research. These are discussed below and at the end of each paragraph a key question is proposed, emphasising a crucial issue.

Perceived issues in outdoor education

1. There appears to be a wide variation in the practice standards in outdoor education both in New Zealand and internationally. A rock climbing example highlights this. Climbing awards exist for the American Mountain Guides Association (AMGA), Australian Climbing Instructors' Association (ACIA), Association of Canadian Mountain Guides (ACMG), and British Mountaineering Council (BMC). Each of these organisations has determined a standard for rock climbing instruction. While there is some common ground, there is little consistency and the standards appear to be arbitrarily determined by different groups of experts. For example, the New Zealand Outdoor Instructor Association (NZOIA) rock climbing scheme has been criticised by some as setting too high a standard and by others for not being rigorous enough (Moyle, 2004).
   How are standards set for assessment?

2. Assessments of live performances or technical practical skills are often ephemeral, providing little evidence and creating an issue of sufficiency for teachers. Furthermore, students are often assessed in a limited range of venues and on a limited range of practices. It is difficult to assess students in multiple areas of capability. Teachers instead infer capability based on small samples of evidence. Teachers, using their professional judgement, presume that students can or are likely to be able to perform and apply knowledge to non-assessed contexts or practices. This raises questions around the reliability and validity of the resultant judgement.
   How are teachers addressing issues of sufficiency in assessments?

3. While there is considerable literature around assessment in general, there is little known specifically about assessment processes (Baume, Yorke, & Coggey, 2004; Cross, Hicks, & Barwell, 2001). If assessment judgement leads to decisions about

the awarding of credit for a qualification, *how then are teachers making decisions about student competence? How are teachers ensuring their assessment tasks and practices provide validity and reliability?*

4. Outdoor education is interpreted in a multitude of ways:
    a. skill development for lifestyle sports; enjoyment, relaxation and replenishing the soul (SPARC, 2009a);
    b. intra and interpersonal development (Priest & Gass, 1997);
    c. environmental and sustainability education (Law & McConnell, 2007).

   Therefore, the intended purpose or expected outcomes of each programme can be different. This has resulted in outdoor education programmes offering differently focused curricula, which can create tension between employers, practitioners, academics, and students. Depending upon one's interpretation of outdoor education, the competence of any teacher is likely to be understood differently. There is also concern regarding the definition of competence of outdoor teachers. Competence in an outdoor education context has often been perceived as the possession of skills and abilities in pursuits. This possibly neglects hard to assess aspects of competence, such as leadership, attitudes, empathy, reasoning, judgement, problem solving, and depth of understanding, (Lugg, 2003). Assessments have traditionally focused on the more readily observable and measurable dimensions of specific technical skills, performance and factual knowledge. This highlights a challenge faced in education when teachers ask how they should assess the less tangible aspects of good teaching or learning. While answers may lie in communities of professional practice (Wenger, 1998), questions like this are difficult to answer. *How do teachers define and practice outdoor education? What are teachers actually assessing?*

5. Finally, various authors contend that assessment influences learning (Biggs, 2003; Ramsden, 1992). According to Biggs (2003) students tend to learn what they think is going to be assessed. The better the assessment aligns to the learning outcomes, the better the learning will reflect the outcomes and purpose of the course. *What then are the practices of teachers towards alignment with stated objectives and promoting learning?*

The issues, questions and commentary highlighted above illustrate the complexity and uncertainty of assessment in outdoor education and possibly assessment in any educational field. Because of these ongoing issues, the following research question and sub-questions directed this research study.

The Research Question

The question that guides this study is:

What are the assessment practices of tertiary teachers delivering outdoor education qualifications throughout New Zealand?

Sub-questions

1. *What are the assumptions underpinning teacher assessment practices?*
2. *What assessment processes do they use and why do they use them?*
3. *What values are teachers communicating through their assessment practices?*

Rationale for and significance of the research study

Assessment can be seen as a key driver of student learning, since it appears to affect both how and what students study. Ramsden (1992) points out that students often see learning as what they have to do to pass, and Biggs (2003) calls this 'backwash', where the assessment influences what and how students learn. Furthermore, students, parents, employers and governments want assessments to be valid, reliable and fair (S. Brown & Glasner, 1999). If assessment is as influential as suggested, educators need to understand the effects of their teaching and assessment practices, but there has been extremely limited research on New Zealand outdoor education, and no known research that examines the assessment practices of outdoor education teachers in higher education. This study attempts to identify, document and clarify current assessment practices and to analyse them against current educational thinking on assessment. The evidence should contribute to a clearer understanding of current practices in outdoor education in the tertiary sector. It may also assist student learning by suggesting possible future policy and advancing assessment practice in outdoor education. The study could also be a base for further research in the area.

Chapter Two: Context of research, supported by a review of the literature on contextual issues

Setting the scene

Although outdoor education is a relatively new kind of learning within formal education, it has a long and rich history. Its roots can be traced back to the Greek philosophers, and in more recent times to Kurt Hahn's influence through developing Outward Bound in the 1940s (Stetson, n.d.). Within New Zealand, outdoor education reaches back to the late 1930s when New Zealand teachers began to emulate their British counterparts in taking children camping. However, it was not until late 1999 that outdoor education became a part of the formal school curriculum in New Zealand, through the health and physical education curriculum. Previously, it had been an activity outside the formal curriculum (Lynch, 2006).

Outdoor education arose as an innovation to the standard and formal school curriculum, a practice which is largely based on a progressive liberal ideology (Dewey, 1938/1997). This sees schooling as a process involving the whole child, not just the mind (Lynch, 2006). Its success has been through providing an alternative to didactic methods by using real life examples in the learning process (Hovelynck, 2001).

Defining Outdoor Education

Brookes (1991) proposes that outdoor education is relative to time and place and that defining outdoor education can neglect changes across time, space and culture. Yet if there is not some form of collective understanding it is difficult to further the theory and practice of outdoor education. Neill (n.d.) offers over twenty different definitions, which represent somewhat of a timeline or evolution of outdoor education. The main outcome of this is it provides a staging point to discuss outdoor education. He identifies two key themes: a psychosocial focus and an environmental focus. Priest and Gass (1997) contend that the psychosocial practitioner is concerned about the development of people through direct and purposeful experiences that appropriately challenge learners and which involve natural consequences. Law and

McConnell (2007) define the environmental practitioner as having "a multi-disciplinary approach to learning that develops the knowledge, awareness, attitudes, values and skills that will enable individuals and the community to contribute towards maintaining and improving the quality of the environment" (p. 9).

In New Zealand, outdoor education has become commonly classified into four main categories (Lynch, 1993).

Outdoor Recreation Education – Outdoor recreation education as outdoor education has generally manifested itself through outdoor pursuit experiences, for example, rock climbing, whitewater kayaking, and mountaineering. These can be classified as lifestyle sports (Tomlinson, Ravenscroft, Wheaton, & Gilchrist, 2005), where individuals tackle 'challenges' which deliberately introduce risk through natural or artificial/architectural environments, so the performer can utilise their skills to overcome the risk of injury or death (Wheaton, 2004). These sports/activities/experiences provide participants with the opportunity to escape from everyday mundane routine and allow participants to feel a sense of well-being, accomplishment and success (Mortlock, 1984).

Education Outside the Classroom (EOTC) – EOTC is a generic term describing any curriculum-based school activity that takes place outside the four walls of the classroom, ranging from a museum or marae visit, sports trip, outdoor education camp, to a rocky shore field trip. Its practice is based on having students engage in learning opportunities beyond the classroom, that relate their classroom learning to real contexts in the wider world. It aims to create student curiosity in order to stimulate and promote opportunities for further learning (Ministry of Education, 2009)

Adventure Education – Adventure education aims at the development of people, groups, and teams through direct hands-on outdoor experiences, in order to develop interpersonal and intrapersonal skills. This usually occurs through outdoor recreation – challenge activities, outdoor pursuits and problem based learning. Understood as activities:

> Where the individual experiences a state of disequilibrium by being placed in a novel setting (placement in an environment that one is not familiar with) and a

> cooperative environment while being presented with unique challenges in situations that lead to feelings of accomplishment which are increased by processing the experience which promotes generalisation and transfer to future endeavours (Luckner & Nadler, 1997, p. 257).

Environmental / Sustainability Education – refers to education about ecosystems, resource use and misuse, and human impact and relationships with the environment. It consists of education for and about all environments, not just natural environments. This can also be further divided into two focuses, stewardship, and activism (Disraeli, 2006):

> **Stewardship** focuses on helping students to develop an awareness of issues relating to the protection of vulnerable habitats through an understanding of, and hopefully a connection to, natural environments.
>
> **Activism** is educating for the environment through personal action, social change and policy development, which supports sustainable living practices.

For many, outdoor education is an umbrella term that encompasses interdisciplinary curricula. It is at this junction that academics and practitioners begin to disagree. Haddock (1999) points out that the term "outdoor education" is often used interchangeably with "outdoor recreation". This tends to muddle the differences that exist in the various approaches and intended goals, leading to misunderstanding when used generically. Academics such as Brown (2008) and Irwin (2008) argue that outdoor recreation is not outdoor education. Instead, they suggest outdoor education places a stronger value on environmental and sustainability education. Brown (2008) introduces the idea of 'place based' education, where outdoor education gives greater emphasis and acknowledgement to 'place (s)' to help students make sense of their personal and communal identity. This challenges the conception that outdoor education requires high impact adventurous activities. EOTC tends to be less controversial as a classification.

While the four classifications provide some clarity, and create a sense of where their boundaries may lie, the practice of outdoor education usually produces overlaps with each or part of the other classifications. For example, where students might be learning outdoor recreation skills, they can also be learning about themselves and the

environment. It is therefore important to recognise that outdoor education is a diverse practice with various outcomes and purposes.

Outdoor Education in New Zealand

Lynch (2006) argues the demise of outdoor education occurred when it changed from informal extracurricular activity and became formalised within mainstream school education. The academic demands created by formal qualifications reduced outdoor education, she says, to a collection of isolated performance standards within many New Zealand schools. While outdoor education may have gained formal recognition and status within New Zealand's education system, teachers argued that this was nonsensical as outdoor education was a practice based on the particular needs of their students, not a subject in a formal curriculum. Furthermore she argues "there was a strong sentiment among many outdoor education advocates that if a specific syllabus was produced, outdoor education would become a formal subject and thus examinable but examination of predetermined material runs counter to the experiential ideology to which many advocates adhered" (p. 155). Hovelynck (2001) suggests that this is because it shifts teaching back towards didactic methods and thinking, where outdoor education is influenced by teaching practice that can be overly assessment focused, causing the learning alternative to be lost.

Physical activities have played, and continue to play, a central role in outdoor education. Assessments in secondary schools and polytechnics have tended to focus on technical pursuit competence at the expense of other competences (Irwin, 2008). This is perhaps due to four main influences: the qualifications and unit standards of Skills Active (the industry training organisation); New Zealand Outdoor Instructors Association's (NZOIA) professional qualifications; the way curriculum has been developed in many polytechnics, where industry has influenced each local polytechnic's curriculum based on local vocational needs (Irwin, 2008); and the way teachers in general have framed outdoor education. According to Irwin (2008), it is evident within New Zealand's outdoor education programmes that there is a strong focus on assessing knowledge of and skills in technical outdoor pursuits. The result is that outdoor education in schools and polytechnics has ended up being positioned

in the realm of outdoor recreation education, where the acquisition of technical pursuit skills are highly valued.

Jones (2005) argues this is may be a valid approach as part of a higher education programme where a student begins a career in outdoor education, but it should not be an end in itself. While technical skills are key aspects of being competent, their position within outdoor education should not displace other skills and knowledge. Jones says, technical skills provide a platform from which other skills can develop. Technical competence therefore maybe seen as a foundational aspect of safe outdoor education practice. Outdoor education teachers must be confident, comfortable and competent in their working environments before they can offer programmes in outdoor education with students.

Nonetheless, both Irwin (2008) and Martin (2005a) argue that outdoor education has an unsettled relationship with outdoor recreation and adventure education. This is clearly affected by its place within New Zealand's secondary school health and physical education curriculum (Ministry of Education, 1999), and also by the way the public generally associates outdoor education with physical education, outdoor pursuits, challenges and adventure in the outdoors. Irwin (2008) asserts that adventurous pursuits or challenging activities place outdoor education in a juxtaposed position to outdoor recreation and adventure education. Martin (2005a) further stresses this by claiming the goal of developing human potential through challenging adventurous pursuits "acts to obliterate the difficulties identified in creating a sense of belonging with and learning about nature within this context" (p. 49). This raises concern over the place of outdoor recreation and adventure education within an outdoor education context.

In contrast, according to Ward (2007) outdoor recreation can occur at many levels, from enjoying autumn colours in a local park to extreme mountain climbing adventures. Its purpose and benefits can vary, providing:

- enjoyment, physical and mental wellbeing as a result of being active;
- bonding with family, friends, and cultural identity;
- learning benefits – environmental and conservation awareness, and personal development;

- economic/employment opportunities.

The Outdoor Recreation Strategy (SPARC, 2009a) identifies that

> New Zealand's heritage has, to a large extent, been shaped by the special features of our unique environment, outdoor lifestyles and adventurous spirits. SPARC (2009a) contends that landscapes and activities define our communities and quality of life; they also underpin New Zealand's tourism industry and innovative outdoor equipment and clothing design and manufacture industry (p. 2).

The Outdoor Recreation Review (which formed the basis for the Outdoor Recreation Strategy) was concerned that participation levels are reducing through population growth, increasingly diverse communities, changing urban structures and growth, societal and lifestyle change, changing and diversified recreation interests and increasing environmental pressures. Lower participation levels are likely to lead to a loss of skills and a decline in understanding of outdoor etiquette and knowledge about the environment that is so vital to outdoor recreation. This suggests that outdoor recreation skills and knowledge, along with a relationship to environments are essential and integrated aspects of being outdoor education, and should not be separated entities as Irwin (2008) suggests.

SPARC (2009a) clearly articulates that to encourage participation:

> learning and development and safety management will need to be enhanced if more people are to be provided with the opportunity to develop skills for participating, leading, gaining greater understanding of the environment and managing risk. For young people these opportunities will come predominantly through outdoor education at school while community-based opportunities such as youth groups will also be important. Skilled instructors and teachers are critical in this process (p. 6).

This highlights the need for instructors and teachers to be skilled in outdoor recreation and have the skills, knowledge and ability to help people build a connection to and with environments. Environmental education and outdoor recreation are interconnected through sharing the outdoor space they utilise. They, do though, remain separated by their purpose, the choice of activities used and the way they are used.

In higher education, Martin (2005b) and Munge (2007) both argue the move to mainstream education was, in part, driven by perceptions that there was a need for a broader knowledge base than has been provided by the apprenticeship style of training, and by outdoor educationalists seizing an opportunity to provide an alternative pathway into the industry. However, the problem of what constitutes a fit-for-purpose outdoor education qualification/programme remains a current dilemma for teachers, and there is a tension between employers', teachers' and students' wants and needs. Several authors argue that employers on the one hand want graduates who can enter employment with minimal need for further training and supervision, who are aware of workplace needs and requirements, and preferably offer more than just beginning competence (Barnett, 1994; Chapman, 1999; Martin, 2005b; Munge, 2007). Chapman (1999) suggests tertiary teachers tend to see themselves as preparing graduates with broad generic skills and a grounding in academic learning, where lifelong learning and transferable skills are often valued more highly than specific competences that employers may seek or students want. Students also bring their own expectations. Irwin (2008) says students entering tertiary education tend to hold a perception that outdoor education is mainly about learning through and about outdoor recreation. This is a direct outcome of what and how outdoor education has been taught in most New Zealand secondary schools.

Many outdoor education programmes in New Zealand have aligned themselves to outdoor recreation education within outdoor education, which tends to base the qualifications on a skills competency training and assessment approach. Students tend to seek out programmes which offer qualifications that can be gained quickly as a way to fast track their learning. Wenger (1998) warns this practice may actually deny the student a fulfilling and confident role in a community of practice. Dumble (2002) supports this notion. He comments, that the shorter the course, the less likely the student is to be able to demonstrate sound judgement and decision making for no other reason than they will not have appropriate time to gain sufficient experience in an instructional or teaching setting. Education in this form is potentially lethal in that it may produce graduates who are highly skilled technically but who have minimal experience – a combination that perhaps has led to some of New Zealand's recent outdoor tragic accidents (this is discussed later).

Furthermore, it is suggested that this approach has led to practices in outdoor education which value and support male paradigms, where competition, skill acquisition and mastery of pursuits (Irwin, 2008) are more highly valued than interpersonal communication, people skills and the general ability to work with others (Lugg, 2003). Assessment practices are said to be reflective of teachers and institution values: they reveal what is paid attention to and therefore what is valued as important (Astin, 1993). This raises questions for teachers to consider, especially if tertiary level outdoor education should lead the field in challenging oppressive forms of practice (Allin, 2000; Carter, 2000; Humberstone, 2000).

Within a tertiary outdoor education programme, there remains an emphasis on physical and technical performance (hard skills), rather than, for example, interpersonal and facilitation abilities (soft skills). Kiewa (2001), points out that even though lecturers make deliberate efforts to adopt an inclusive approach, a culture of 'hard and fast' is the norm and this is reinforced by assessment of technical skills at the expense of other facets of practice. Lugg (2003) supports this, but suggests this does not mean that technical skills should be valued less – rather an equal emphasis should be placed on interpersonal and process skills for reasons of gender equity and to help redefine outdoor education. Otherwise, as Allin (2000) Carter (2000) and Humberstone (2000) say, outdoor education remains locked into traditional male paradigms, which Green (1994) describes as where a normative emphasis is placed on physical strength, speed and technical expertise. Employers echo these interpretations: a skilled outdoor educator is one who has a range of effective interpersonal skills to support technical capabilities (Carter, 2000; Neill, 1997).

This raises questions as to what constitutes competence and who decides. Failure to reconsider outdoor education for today's world ignores changes to the social and environmental landscape which women are a part of and where issues of climate and ecology are being brought to the forefront (Brookes, 2003; Hovelynck, 2001; Irwin, 2008; Martin, 2005a). Outdoor education teachers need to remember that outdoor education is no longer a part of an imperialistic and militaristic world in which students are being trained for war (Beedie, 1995/6).

While the practice of focusing assessment on technical skills is historical, and justified in terms of enhancing safety, the practice of avoiding soft skills assessment perhaps offers an insight into the challenges of teaching interpersonal and facilitation skills, judgement, decision making and the ability to assess risk and challenge levels for educational purposes. Soft skill acquisition may well be desirable, but is often missing from descriptions of course outcomes and is perhaps not assessed due to the complexity of assessing them in short periods of one-off assessments. Often, hard skills are taught and assessed through a reductionist approach, a practice that has broken down learning[1] and assessment[2] into small elements. The same approach is not applied for the assessment of soft skills, perhaps due to the perception that these skills are subjective. Assessing outdoor education through a reductionist approach encourages the view that acquiring professional and pedagogical knowledge and expertise is reducible to a set of separate or separable standards (Yandell & Turvey, 2007). Such assessment methodology also neglects the idea that competence-based assessment can allow for unpredictability. Accordingly, Barnett (1994) says it is ultimately incoherent to write a standard to assess the unexpected. Reducing complex segments of skills into small assessment tasks is challenged by Biggs (2003), who argues that isolating segments of practice, rating each independently and then assimilating them together does not provide an assessment which can be considered valid[3].

Clearly, outdoor recreation education as a part of outdoor education creates tension for teachers, especially in higher education programmes. The issue that Irwin (2008) and Martin (2005a) discuss is perhaps due to outdoor education type programmes being initiated from their inception in the context of outdoor recreation rather than

---

[1] **Learning**
The process of learning is defined as acquiring new or modifying existing knowledge or skills through experience, practice, or study, or by being taught.

[2] **Assessment**
The process of assessment is the gathering, analysing, interpreting and using information about students' progress and achievement to improve teaching and learning.

[3] **Validity or a valid assessment**
Validity in assessment is defined as an assessment that measures what is intended to be measured. This is judged by how well the assessment matches the published learning outcomes for each course.

educational interests. Outdoor education perhaps needs to be viewed as more than a matter of leading groups in outdoor recreation experiences, where risk and challenge are high motivators, although, the Outdoor Recreation Strategy predicts a general lack of competence in the outdoors if the appropriate recreation skills are not taught. While academics may be promoting the idea of outdoor education being used for sustainability and environmental education, and encouraging a move away from the exhilaration of outdoor recreation education and experiences, the demand for such experiences remains. The challenge is perhaps how to integrate sustainability and environmental education into a programme, which meets the demand for outdoor recreation skills.

Recent changes to the outdoor education landscape

How outdoor teachers are educated, trained, and assessed has changed over the last two decades. The key influencing factors are the introduction of professional standard bodies (predominantly NZOIA), outdoor education finding a place in tertiary education, and new legislation (Industry Training Organisation Act 1992 and the Health and Safety in Employment Act 1992 (HSEA)). The section below briefly introduces these influences.

Legislation

The Health and Safety and Employment Act (Health and Safety in Employment Act, 1992) requires employers to take all practicable steps to protect their workers and those in their care. As previously discussed, outdoor education can involve exposing teachers and students to hazards and risks that can result in serious harm. What distinguishes outdoor education from other classroom education and other industries is the deliberate use of risk with students in an educational setting. It is therefore common practice and required by the HSEA to operate under some form of hazard identification and safety management system. The management of risk, is in theory, a systematic application of management policies, standards and procedures to the task of identifying, analysing, assessing, treating and monitoring risk (Brown, 1999 as cited in Zink & Leberman, 2001). This implies that with the right management systems, risk can be reduced to acceptable levels. However, what is ‘acceptable’ is

subjective and debatable. The management of risk is not the same as the removal of risk. Risk remains, but it is keep at an appropriate level to ensure the safety of students. This has caused extended and ongoing discussions for teachers around appropriate levels of risk and strategies for the reduction and management of harm in outdoor education environments, and clearly, this has not stopped incidents and serious accidents occurring.

As a result, the Health and Safety in Employment Act (Health and Safety in Employment Act, 1992), has affected outdoor education practices in both a positive and negative manner. It is difficult to argue with the fundamental purpose of HSEA, that of keeping people safe and healthy while they work. This has resulted in the development of Standard Operating Procedures and Safety Management Plans. The potential problem with set ways of operating is it can lead to a 'rule book' culture. It creates a culture where professional judgement and experience is less valued, and written documents and plans are given higher status. It produces an environment where following and remembering all the rules is considered imperative, yet at the same time almost impossible, due in many cases to excessive volumes of rules, regulations and procedures.

Standard operating procedures (SOPs) and safety management plans (SMPs) are common in most organisations involved in outdoor education in New Zealand (Haddock, 2004). Following specified training and practices in theory leads to teachers knowing what to do in normal and emergency situations. However, Austin (2005) argues not all situations are 'standard' and this approach of following rules does not allow for novel and unfamiliar situations. This leaves teachers in a double bind. If they do not follow the rules they lay themselves open for blame (especially if there is an accident), and if they do follow the rules, these procedures might not fit the situation and could still compromise the safety of their students. This creates tension between the 'training package', doing things in a set way and the teacher working using professional judgement and integrating a range of skills and knowledge at one time to meet the demands and management of risk in an adventure context. Following procedures may be appropriate in many situations. However, does such an approach fit an adventure context that is highly dynamic and potentially dangerous?

It is a delicate and subtle balance between an approach which says you must do as your management has trained and directed you to do and an approach where teachers are trusted to do the right thing for their students in their care. It is a continuum where the HSEA has potentially forced the pendulum to swing too far to one side – towards a 'systems based' approach. Too much of either perspective can compromise safety and the benefits of challenge based learning. It is difficult to argue against the HSEA – that keeping all people safe is sensible. However, if we look at this critically, the way the Act has been interpreted has left us with a legacy where what is considered 'good practice' emphasises written plans in favour of professional judgement and good in-the-field decision making. Professional practice clearly becomes diminished in favour of these plans and the question arises whether a risk adverse society and the rule book culture of the HSEA approach has in part contributed to recent accidents. This poses a problem for tertiary teachers: what should they be teaching and assessing future graduates for if this is the culture of our society?

Industry Training Act (ITO) 1992

This Act provides for the recognition and funding of organisations (known as industry training organisations) setting skill standards for, and administering the delivery of, industry based training. It provides an avenue for the skills and knowledge in outdoor recreation education and adventure education to be formally recognised through the formation of Skills Active (the main ITO for this industry). The Act requires the ITO to identify activities for funding. These activities include the development of standards, training pathways, assessment, leadership, and qualifications for each sector of industry. Skills Active has generally been unsuccessful in adequately providing these functions in the outdoor recreation education field (Paul Richards, personal communication, Skills Active employee, Dec. 2010) as these qualifications have not been embraced by industry and have mostly been used in secondary school outdoor education programmes.

Key organisations within New Zealand

Skills Active the industry training organisation (ITO) responsible for the sport, fitness, community recreation, snow sport, outdoor recreation and Nga Mahi a te Rehia sectors. Outdoor education fits within this ITO through the overlap of the activities used.

Aviation, Tourism, Travel Training Organisation (ATTTO) is the industry training organisation responsible for the aviation, travel, tourism and museum industries. This organisation overlaps with Skills Active by providing qualifications in the tourism area for outdoor guides who provide outdoor educational experiences for tourists, through the likes of sea-kayaking and horse trekking.

Education Outdoors New Zealand (EONZ) is a national professional association of outdoor teachers who aim to promote quality outdoor education. Outdoor education is seen to encompass opportunities for learning: **in** the outdoors, **about** the outdoors, **through** the outdoors, and **for** the environment, and as such encompasses all areas of the school curriculum. This organisation acts predominately as a voice for teachers in schools.

New Zealand Outdoor Instructors Association (NZOIA) is the professional body for professional instructors who aim to provide a nationally recognised assessment scheme for outdoor instructors (including teachers) who use outdoor pursuits in outdoor recreation education and adventure education. The scheme is performance based, reflecting the minimum requirements for a person to lead or instruct others in a given environment and situation. Presently NZOIA administers a nationally recognised award scheme in bush, alpine, abseiling, rock climbing, canoeing, kayaking, sea kayaking and caving. In recent years, they have also been offering training workshops. They offer a voice for the outdoor instructor who is not in formal school outdoor education.

Maritime New Zealand (MNZ) is connected to the outdoor education sector by rule 80 of the Maritime Safety Act, which requires commercial raft guides to be qualified. The rest of the outdoor sector is under no other legislative obligations that require instructor/teachers to have certain qualifications to run or lead an outdoor activity.

This requirement may change as the government are currently reviewing safety in adventure and outdoor activities (Department of Labour, 2010). Rafting has been an activity used in outdoor education. However, it is normally associated with adventure tourism.

New Zealand Mountain Safety Council (NZMSC) is a national organisation promoting safety in land based outdoor activities. It facilitates the setting of standards, offers training, distributes resources and advocates publicly so that more people can discover and enjoy New Zealand's outdoors safely. It offers training and assessment for volunteer instructors in: abseiling, avalanche awareness and backcountry avalanche, bush craft, firearm safety, outdoor first aid, outdoor leadership, risk management and river safety.

Water Safety New Zealand (WSNZ) is the national organisation responsible for ensuring all New Zealanders play safely in the water – whether at home, at the pool, the beach, in lakes and rivers or out at sea. WSNZ sees educating the public on water safety as one of its fundamental tasks as education is proven as the most effective and efficient way to promote and improve water safety. They do not offer qualifications, but they fund organisations such as NZOIA to help deliver the water safety education and skills message.

New Zealand Mountain Guides Association (NZMGA) was established to provide a consistent training programme for professional Mountain Guides in New Zealand. It is a member of an international federation of 26 mountain nations, which share their skills and knowledge and set standards. They offer 10 different qualifications and set and maintain standards for admission to, and practice of, the mountain guiding and heliskiing profession.

Sport and Recreation New Zealand (SPARC) facilitated the development of the Outdoor Recreation Strategy for New Zealand until 2015 (SPARC, 2009a) in conjunction with the Department of Conservation. The purpose of the strategy is to get more people participating in outdoor recreation so that the many benefits associated with outdoor recreation can be enjoyed by individuals, communities and the nation (SPARC, 2009a). Although SPARC is not directly involved in assessment,

its current outdoor recreation strategy may influence the direction of outdoor education. It is involved in everything from supporting elite athletes to getting out into local communities and encouraging people to become active. SPARC funds the Sir Edmund Hillary Outdoor Recreation Council to advise on strategic investment decisions relating to outdoor recreation. One of the key findings of the SPARC outdoor recreation review conducted over 2007/08 was that nearly three quarters of New Zealanders want to increase their overall participation in outdoor recreation. Among the council's roles is to provide leadership and advice on:

- Promoting participation in outdoor recreation;
- Developing greater capability in outdoor recreation;
- Decisions on investment in outdoor recreation organisations;
- Developing a strategic sector-wide approach for outdoor recreation;
- Fostering improved collaboration and co-ordinated planning within and between the outdoor recreation sector and other relevant sectors;
- Raising the profile of the positive values and impact of outdoor recreation (SPARC, 2009b).

Tertiary providers, predominantly the polytechnic sector, provide outdoor education, outdoor recreation and adventure tourism qualifications at levels 3-7. Programmes and courses have evolved from local industry consultation and, as such, are local provider qualifications, often with embedded NZOIA and other organisation qualifications such as St John First Aid.

As can be seen, New Zealand has a plethora of organisations involved in outdoor education. There are varied outdoor education qualifications, education, training, and assessment approaches and pathways that reveal a somewhat disjointed industry, one which involves 'turf wars', duplication of resources and a range of differing standards. Adams (2007) states "the New Zealand outdoor sector suffers from a lack of structure and weak government support. These are fundamental causal factors in the confusion that exists around outdoor standards and qualifications in New Zealand" (p. 3). While Skills Active and ATTTO are mandated by legislation, and therefore directly responsible for qualifications within outdoor recreation education, adventure education and adventure tourism, they have been unable to develop qualifications which are respected and valued by the industry. Polytechnics have

instead offered local qualifications, utilising small national awards embedded within these qualifications. They have provided the bulk of education and training in this field since the early 1990s. Private Training Establishments (PTEs) have offered the ITOs qualifications (unit standards) in order to gain funding, even though these qualifications have had little acceptance within the outdoor education sector, leaving employers with uncertainty about the competence of graduates from these programmes (SPARC, 2007). The professional standards body NZOIA has gained the greatest traction and become the accepted standard for many throughout New Zealand. However, few school teachers have committed to these qualifications due to their cost, the level of standards set, the general lack of funding for teachers to achieve these qualifications, and the narrow focus of the qualifications, which are perceived as putting too much emphasis on technical performance.

Because of all this, these organisations have tended to work in opposition to each other in order to protect their self-interest. However, late in December 2009, a significant memorandum of understanding was achieved between Skills Active, NZOIA and MSC (Carpenter, Gunn, & Cant, 2009), and further talks are under way with NZMGA. While this is a major step for the outdoor recreation education and adventure education sector in terms of collaboration, especially in the areas of national standards, qualifications and assessment, the challenge is how these organisations will develop qualifications that reflect the broad aspects of outdoor education and provide assessments which embrace the holistic value and the complexity of outdoor education.

Conclusion

What emerges from the above discussions is that outdoor education is a contested space, one where old and new paradigms meet. The possible use of risk, adventure and outdoor recreation experiences are questioned. The issue of what outdoor education aligns itself with is paramount. This affects the resulting curriculum design, teaching practices, content, and the assessment type and practices. There is little doubt that outdoor education purpose, principles, pedagogy and assessment are currently being challenged. The differing expectations converge into a number of conflicting perceptions: the purpose and the function of higher education and how

this affects and relates to outdoor education, the common and emerging understandings of outdoor education programmes, and assessment practices in outdoor education, such as how student capability is evaluated (Biggs, 2003; Chapman, 1999; Green, 1994; Lugg, 2003; Martin, 2005b; Seaman & Coppens, 2006). These perceptions do not exist in isolation from one another; they are interrelated and affect each other. Small changes to any one of these perspectives could influence how outdoor education is taught and assessed. The broader perspectives of outdoor education perhaps need to better connect to practice in order to balance Irwin's and Martin's concerns. What becomes apparent from the various perspectives is that outdoor education reflects teachers' values, which become reflected through programme design, what teachers teach, what students are supposed to learn and what is assessed.

This research study positions itself among this political change, focuses on the teachers who are teaching and assessing in the polytechnic sector and seeks to understand their assessment practices.

Chapter Three: Literature Review

In this chapter literature pertaining on assessment was reviewed as it applies to higher education, with particular focus on common practices, issues in and around assessment, and emerging ideas or trends. Assessment as an area of study within education, has a relatively short history (Hill, 1998). Joughin (2009a) warns that a comprehensive review of assessment is, under any circumstance an ambitious undertaking given the vast amount of literature on this matter. To keep this research study tenable and, at the same time, comprehensive, I have reviewed relevant literature from recent years, and followed this with an overview of antecedent literature in order to provide a rich synthesis of themes supported by interspersed critical commentary. This approach is perhaps best compared to a rolling snowball, which gathers more and more snow as it rolls (literature leading to more literature), with chunks falling off (irrelevant literature being left behind). This synthesis of other people's ideas provided me with a platform of understanding that offered new perspectives and broadened my understanding of assessment, to help inform this research.

Black (1998) suggests assessment practices around the world connect to each country's history, the culture of their educational system and broader social influences. Biggs (2003), Boud and Falchikov (2007), Boyes (2004), Joughin (2009b), and Lugg (2003) all claim that the process of assessment is complex and its purpose and design are highly contested and value-laden. The theory and the practice of assessment are widely debated by academics, industry, governments, students and various other stakeholders within society, all of whom have their own agendas, assumptions and perspectives on the matter.

Dunn, Morgan, O'Reilly and Parry (2004) say that assessment is used to accomplish several different purposes. They observe, that while it may seem obvious that the purpose of assessment is to measure student learning this thinking is overly simplistic. Nevertheless, they continue, it remains the dominant perspective. In the opinion of Dunn et al., (2004), the role and purpose of assessment is far broader than that of measuring student learning and maintaining pre-set standards of achievement. They propose that effective assessment is that which diagnoses student difficulties,

measures student achievement (with particular focus on improvement) over time, motivates students, judges mastery of essential skills, evaluates teaching efficacy, and provides feedback to students. Assessment can clearly encompass various purposes, but some are contradictory in terms of approach and anticipated outcome. As Earl (2003a) says, it is not possible to use one assessment process for the many purposes that various stakeholders want them to fulfil. This situation creates tension with respect to assessment implementation because the various stakeholders involved act according to their various perspectives on what assessment is and should achieve. Angelo (2003), for example, argues that assessment directed at assessing extent and nature of learning simultaneously meets other, often covert, needs that in practice are difficult to marry. Earl (2003b) brings clarity to Angelo's claim when she points out that assessment in many societies acts as a way of selecting people for entry to higher education and for privileged professions, although the espoused purpose is to measure students' achievement at the end of a course of study. Broadfoot (1996) summarises this dichotomy well by asking whether assessment is about our ability to measure (and sort) or to learn. Gipps (1994) proposes that assessment design should come into play after we have decided our purpose.

Earl (2003b) highlights the view most of the twentieth century learning was conducted according to a behaviourist model that focused on acquiring specific, discrete skills and facts in a hierarchical sequence that signified teachers as transmitters of information, in an environment where education was very content focused. As Hill (1998) points out, learning is a complex process, which means that the assessment purpose must be closely linked to models of learning. Biggs (2003) and Black (1998) believe learners construct their own understanding of the world they live in. Transmission and acquisition of new knowledge therefore has to be moulded to fit each learner's own schema. Biggs (2003) emphasises it is what students have to do in order to gain knowledge and understanding that matters. Learning he says, "is constructed as a result of the learner's activities" (p. 11). Turning his attention to what this thinking means for teaching, he argues that "a good teaching system aligns teaching method and assessment to the learning activities stated as objectives, so that all aspects of the system act in accord to support appropriate learning, called constructive alignment" (p. 11).

Biggs (2003), Marton and Saljo (1976) and Ramsden (1992) describe the way students engage with learning as being deep or surface level. Surface learning can be understood as a student simply wanting to get the learning and assessment task completed and out of the way with the minimum of work while appearing to meet the course requirements. Deep learning, in contrast, involves students engaging with learning and assessment opportunities with a genuine desire to complete the task as well as possible and in a meaningful way in order to learn as much as possible. Atherton (2009) suggests this may correlate fairly closely with motivation: "deep" with intrinsic motivation and "surface" with extrinsic. Although he suggests that they are not necessarily the same thing, either approach can be adopted by a person with either motivation.

Biggs (2003) proposed that this probably means that if a student thinks an assessment task requires memorising and recall then the student is more likely to learn in a way that matches the level of expectation and demand, resulting in a relatively low level of cognitive engagement leading to surface level learning. On the other hand, if a student thinks a learning/assessment task requires analysis, critical thinking, and the application of basic knowledge, then the student is more likely to adopt a learning strategy that encourages higher levels of cognitive engagement and results in deeper learning. Therefore "what and how students learn depends to a major extent how they think they will be assessed" (Biggs, 2003, p. 140). In practice, however, the scheduling of assessments can create high peaks of workload for learners where they have limited time to spend on any one piece of work. This, in effect, can limit the possibility of in depth learning, as there is simply insufficient time. Teachers perhaps need to consider their whole programme to ensure the in-depth learning they desire and promote is actually possible.

Dunn et al., (2004) suggest assessment should lead students to learn the fundamentals of their discipline area, for example, the basic principles, facts, terminology, concepts, procedures and rules, while also developing their knowledge of their discipline into understanding. Meaning, students should apply what they have learnt to problems, make decisions, and develop meaningful relationships with basic knowledge to solve complex issues. While Anderson and Krathwohl (2001), Biggs (2003), and Bloom (1956) work on educational taxonomies advocating greater

assessment sophistication, many assessments they say remain focused on testing basic knowledge and comprehension. They argue, assessments continue to ignore the challenge of assessing complex aspects of learning such as decision-making, the use of judgement and the ability to apply knowledge to real problems in real time. Bryan and Clegg (2006) add to this by suggesting that teachers are overly concerned with assessing current knowledge instead of assessing generic skills such as the ability to learn. Boud and Falchikov (2007) offer support for this by contending that higher education should be broader and assess students' ability to learn, to solve problems, and to illustrate that they are capable of continuing to learn after graduating.

Boud (2000) also argues that summative assessment has dominated thinking and practice in higher education, in public policy debates and it takes up too high a proportion of staff time, energy and resources at the expense of preparing effective learners. Assessment he says has typically been an act performed on the learner, if learning is the priority of education, then an effective way to change student learning is to change the form of assessment. Boud (2000), also pointed out that summative assessment, ironically, drives out learning at the same time it seeks to measure it. Accordingly, as Boud and Falchikov (2007) and Light, Cox and Calkins (2009) point out that regardless of the flaws of traditional end of course summative assessments they still prevail throughout many higher education programmes.

These practices according to Schwartz (2002) are predominately based on testing recall of basic information, leaving learners with inadequate opportunities to illustrate what they know, understand or can do. It simply provides a way of measuring and ranking a large number of learners against each other – known as norm referencing. Unfortunately, as Dochy (2009) says, students have been seen as subjects who must be tested. Boud (2000) points out that testing students without standards or criterion-referenced assessment, means learners cannot know whether their achievements are a result of meeting an acceptable standard or simply doing better than other students in the same cohort.

Furthermore students' efforts may be geared towards learning only that which they consider the important aspects; aspects which will be assessed (Boud & Falchikov, 2007; Joughin, 2009a; Watkins, Dahlin, & Ekholm, 2005). Biggs (2003) has well

documented that learners quickly figure out what is really required to pass a course of study and become strategic. This has become known as the backwash effect of assessment (Biggs, 2003). Sadler (2007) suggests that educators in general are working with conceptions of both learning and certification of the learning. He suggests higher education needs to prepare students not just for the programme itself, but also to operate in a society in the future.

Summative assessment or assessment *of* learning is a formal process used to see if a student has acquired the skills, knowledge, and behaviours that the course set out to provide. Formative assessment or assessment *for* learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there (Assessment Reform Group, 2002). Black (1998) states that while assessment can be formative or summative, "the traditional dominance of summative function means that formative assessment struggles for its status and development" (p. 120).

Biggs (2003) argues the 'teaching-assessment relation' is crucial for improving learning outcomes and recognised the importance of aligning the assessment system with the course and the instructional approach. Assessment *for* learning can be considered in two ways:

1. Assessment *for* learning according to Black, Harrison, Lee, Marshall and Wiliam, (2003) is where "formative assessment is a process, one in which information about learning is evoked and then used to improve the teaching and learning activities in which teachers and students are engaged" (p. 122). Ecclestone (2007b) says it has a "strong emphasis towards assessment as raising standards of achievement for everyone through better feedback, sharing outcomes and criteria with students, recording achievements and progress" (p. 42).
2. Assessment *for* learning can as Murphy (2006) says, also be understood as assessment which by its design promotes learning and could be used for either formative or summative purposes. This perspective he says, creates opportunities for the use of ongoing coursework and continuous assessment where evidence is accumulated over time and used to help learning but may

also be used for summative reporting. This approach lessens the emphasis on end of course assessments.

These definitions allow various sources of assessment evidence to be used for broad purposes but, perhaps more importantly, it allows information gained to affect teaching and learning activities as well as end outcomes. However, Black et al., (2003) and Gipps (1994) caution that using formative assessment for learning and as summative assessment can inhibit the potential learning because the focus can fall back on assessment for grading. Murphy (2006) argues it is imperative that assessment *for* learning focuses on providing constructive information for students so they can influence their future learning, for example, by providing timely feedback.

Hattie (2009) suggests of all the factors that make a difference to student outcomes, the power of feedback is paramount in the pursuit of learning. Feedback is information provided by an agent (for example, teacher, peer, book, parent, or one's own experience) about aspects of one's performance or understanding. Its main purpose according to Hattie (2009) is to reduce discrepancies between current understandings and performance and a learning intention or goal. Torrance (2007) believes providing criteria and feedback is at the heart of the formative process, but suggests this process can weaken learner autonomy. In addition, Hume and Coll (2009) suggest recent research highlights "teachers are implementing a narrow interpretation of formative assessment in classrooms using techniques that focus on assessment procedures and practices to assure students comply with criteria and achieve awards for external qualifications" (p. 269).

Ecclestone (2007a), Marshall and Drummond (2006) and Sadler (2007) further suggest that teachers have adopted formative assessment strategies that superficially look like assessment *for* learning but are instead employing strategies that lead to instrumental learning. They say, instead of students becoming more self-reliant they receive extensive coaching and practices to meet the demands of a qualification environment. Students actively seek details from teachers about the assessment evidence required, resulting in the learning being narrow, rote, and short-term. Torrance (2007) argues assessment *for* learning has developed in some cases into assessment *as* learning to meet the needs of institutions' compliance and pass rates.

Hume and Coll (2009) suggest this is because, teachers are teaching to the test to ensure better pass rates. Black, McCormick, James and Pedder (2006) agree, saying teachers practices have fallen back into the role of transmitters of knowledge, a negative view of assessment *as* learning. This has lead Torrance (2007) to question this approach as he is concerned that this practice leads teaching and learning environments away from learning and becomes overly focused on assessment procedures and practices.

Assessment *as* learning in a more positive view "emerges from the idea that learning is not just a matter of transferring ideas from someone who is knowledgeable to someone who is not, but an active process of cognitive restructuring that occurs when individuals interact with new ideas" (Earl & Katz, 2006, p. 41). Race (2009) further describes assessment *as* learning where all of the elements of assessment are designed with learning in mind. Volante (2009) suggests that assessment *as* learning promotes student metacognition and DeLuca and Klinger (2010) explain this to mean assessment which focuses heavily on teaching students how to self-monitor so that they can begin to identify learning strategies that best fit their learning needs, fostering self regulated learning behaviour. This they believe better equips students for increased success and promotes life-long learning.

Earl (2003a) says that this enables students to develop the skills needed to be self-motivated and practice using their knowledge to solve problems and make decisions, without waiting for an authority to tell them they have the right answer. In such an environment the learners are actively engaged in the learning and assessing process, and provided with opportunities to take greater control over their learning. As such, they become independent learners and evaluators. Assessment *as* learning therefore involves teachers and learners entering into a negotiated space where what is taught, how it is taught, and how learning is assessed becomes open to collaboration. However, in a climate where governments and other stakeholders demand standards be met, it is not straightforward or even possible to offer control of assessments to a learner. The teacher as an expert is trusted and expected to assess whether a learner has met a standard. This creates tensions, as teachers are forced to meet institutional demands rather than make efforts to develop innovative learning and assessment practices. A culture where pass rates are more valued than learning, understanding,

and capability. It may not be possible to locate assessment totally in the hands of students while acknowledging the legitimate role of certification by teachers and institutions. Negotiating the learning and assessment programme with students may provide better engagement for learning and as a result improve pass rates.

Many of these issues about assessment come from assumptions that it has one main purpose and that assessment is a scientific process that is rigorous, unbiased and mathematically defensible (Dunn, et al., 2004). That paradigm largely views assessment as quantitative or performance as being measured in concrete terms. Woolf (2004) suggests there is wide spread agreement among writers on assessment, that clear assessment criteria are the end result of an effective assessment strategy. Criterion-referenced assessments depend on the assessment tasks being explicit and the performance standards being specific. However, Woolf suggests in practice assessment criteria are often weak. This is because it is difficult, if not impossible, to absolutely specify all of the criteria explicitly and in sufficient detail.

Chapman (1999) argues assessment has always contained a degree of subjectivity, simply because it constitutes a value judgement. Moreover, judgement plays an inevitable role in assessment, not just teacher judgement but also learner judgement. Learning is learning to make judgements (A. Jones, 1999). Yet Woolf (2004) contends assessments against explicit standards or criteria have become widely accepted practice and should have reduced the use of subjective judgement. Jones (2006) suggests this is based on a mythology that caused many stakeholders to believe that assessment criteria could fully and accurately describe learning to enable a skilled practitioner, with limited training, to easily make an assessment decision. However, as Crooks (1988) says, assessment requires teachers to make judgements about what they are marking or observing and then to deliver their judgements for summative or formative purposes.

According to Baumne, Yorke and Coggey (2004) the actual process of assessment has received little attention. They suggest reliability[4] of assessment in higher education has been a problem for a long time. Cross, Hicks and Barwell, (2001) and Chapman (1999) wonder how we are assured of expert judgement where there are concerns about the subjectivity of the teacher. This highlights a perennial problem, one of inter-rater reliability and professional judgement. Mussweiler (2003), however, points out that "human judgement is comparative in nature. When people evaluate a given target, they do not do so in a vacuum. Rather, such evaluations are made within and in relation to a specific context. In fact, any evaluation is relative in that it refers to a comparison of the evaluated target with a pertinent norm or standard" (Mussweiler, 2003, p. 472). Baume et al., (2004) add to this: "In their account the reasons for judgements, sometimes sees assessors iterating between their overall impressionistic or holistic judgements on an outcome and the judgements obtained by strict application of the rules" (p. 459). In other words teachers "in real assessments look even harder for grounds to pass people" (p. 469). Thus, markers are reconceptualising assessment processes as they progress through assessment judgements and creatively apply any rules/criteria. This can, of course, affect the results awarded if markers are changed. Teachers plainly apply "assessment rules on a continuum from 'rules are rules' through to 'whether I like them or not' or 'my judgements are sounder than slavish application of the rules'" (Baume, et al., 2004, p. 470). Crooks (1988) though, points out that one concern about using assessment criteria is that teachers will become straitjacketed into a system that leaves no part for professional, academic judgement in assessing a learner's work or performance.

Professional judgement, according to Dunn et al., (2004) needs a high degree of knowledge, with appropriate and acceptable levels of skills that are backed by benchmarking. McMillan (2000) strongly believes professional judgement is the foundation for assessment. Speck (1998) describes professional judgement as being central to determining grades, but notes that it is often viewed as unpredictable and can appear unreliable. Bruniges (2007) suggests everyday teachers make judgements

[4] **Reliability in assessment or reliable**
The reliability of an assessment refers to the consistency of a measure or a teacher applying consistent judgements. An assessment is considered reliable when the same results occur regardless of when the assessment occurs or who does the scoring.

about the appropriateness of an activity for learning and this is a critical component of teachers work. Smith (1999) says professional judgement is not a simplistic instrument, but an ethical and attentive process that teachers find hard to explain. Professional judgement refers to 'tacit knowledge' applied by skilled professionals. Jones (2006) describes professional judgement as having a strong relationship between 'knowing', the level of expertise, tacit knowledge and that of making sound judgements. According to Tripp (1993) these judgements range from practical, diagnostic, reflective, and critical. He says judgements about learners' achievements and progress is a fundamental skill. Despite all the positive developments in assessment through criterion-referenced assessment, professional judgement still plays a central role.

Crooks (1988) continues by saying that it is more important to assess the right things (validity of the assessment), than ensure the assessment is reliable, as he suggests validity is more important. Once we are assured the right things are being assessed we can seek ways to improve reliability. Accordingly, Crooks suggests that focusing overly on achieving reliability can send assessment practices down a line of defining criteria and applying them rigidly, preventing learners from getting credit for originality or preventing appropriate use of teachers' professional judgement. He realised that teachers tend to treat marks as immutable, not realising that they are subject to both random and systematic errors. For this reason, he believes teachers should exercise professional judgement and make discretionary adjustments if they consider a student deserves to pass or fail. McMillan (2000) supports this view, arguing that although measurement of student performance may seem objective with such practices as rubrics and multiple-choice test, even these approaches are based on professional assumptions and values. Whether that judgement occurs in constructing test questions, scoring essays**,** creating rubrics, grading participation, or combining scores, the essence of the process is making professional interpretations and decisions.

While making a 'judgement call' implies that one is being active, committed, accountable, powerful, and ethical, like a 'courtroom judge' (A. Jones, 2006), teachers live in an ever increasing climate of accountability, where evidence based on clear documentation is regarded more highly than judgement. Professional

judgement in criterion-referenced assessment is intrinsically linked with the notion of professionalism and professional practice (Wenger, 1998). As Wenger says, these should take into account the culture, conventions, traditions, and values of the learning context if it is to be good judgement. It is not a process of isolated and subjective decisions, it is a process performed in a community of experts to report and assist learning.

Conclusion

Assessment remains a highly contested area in education. While there are strong arguments that assessment should support learning, many stakeholders remain anchored to summative high-stake assessments. In general, these assessments still assess a very limited range of knowledge, skills, and behaviours. Criterion-referenced assessment is perhaps more common practice yet it remains engrained in summative purposes, forcing assessment *for* learning and assessment *as* learning into an undervalued position. Assessment *for* learning may have gained some ground in practice, yet governments and employers are still debating policy around measurement of outcomes, again focusing assessment on summative purposes only. Assessment could satisfy stakeholder interest by not only certifying, but also providing feedback to students and helping to equip them for a lifetime of learning. Boud's (1995) comments capture the assessment issue well: "Students can, with difficulty, escape from the effects of poor teaching, they cannot (by definition, if they want to graduate) escape the effects of poor assessment" (p. 35).

Chapter Four: Methodology

In this chapter, the broad philosophical paradigm in which the research approach is located will be identified and the specific methodology, namely illuminative evaluation, will be discussed. Secondly, the research design will be described and elucidated. Finally, the process of data collection, analysis methods, research limitations and ethical considerations will be outlined.

This research is situated predominantly within the qualitative research paradigm. According to Denzin and Lincoln (2008), "qualitative research is situated activity that locates the observer in the world" (p. 4). They say this involves the studied use and collection of a variety of materials that describe routine, problematic moments and meanings in individual lives. It is an approach that attempts to gain in-depth understanding of the phenomenon in question, by capturing the stories of individuals. However, this is seldom possible since individuals hardly ever offer full accounts of their intentions or actions. Consequently, a wide range of interconnected methods is used to gain better understanding. Such an approach is a strategy that adds rigour, breadth, complexity, richness, and depth to the inquiry (Flick, 2002).

Denzin and Lincoln (2008) argue there are no objective observations, only observations socially situated in the worlds of the observer and the observed. Any research is always "filtered through the lenses of language, gender, social class, race and ethnicity" (Denzin & Lincoln, 2008, p. 29). No research methods are immune from prejudice, experimenter bias and human error (Parlett & Hamilton, 1972). The paradigms that surround this research have, therefore, been based around personal epistemological and ontological foundations. We all view life through our own eyes and, as much as I have attempted to cover a range of perspectives, this research will undoubtedly have been affected by my viewpoints, as it is my beliefs that inspired the research in the first place. It is also important to acknowledge that this research has occurred in the light of my own paradigmatic position, as I am also a tenured tertiary outdoor education teacher. To reduce assumptions in this research, a number of external controls were utilised. I sought the use of peer review, as well as seeking critical feedback from my supervisors before and during the gathering of data and prior to this output.

Qualitative research is based on the understanding that there is not one universal truth, people and their experiences are unique and need to be interpreted in context (Gephart, 1999). We all make sense of the world in slightly differing ways, through our experiences, and this influences our choices, resulting in different explanations for similar and different encounters in our lives. Interpretive research considers this uniqueness and values the beliefs and opinions of each person and their situation. There are three main reasons, for a qualitative research approach:

1. The paucity of research in this area, as well as the varying interpretations of Outdoor Education. A qualitative interpretative research approach will reveal individual outdoor education tertiary teacher values along with various and wide-ranging perspectives.
2. An interpretive methodology is appropriate for investigating individual teacher practices in tertiary outdoor education as it allows individual perceptions, feelings, and thoughts to be illuminated.
3. When a new area is being explored a qualitative design is appropriate as it helps to facilitate the acquisition of new insights and understanding. This approach is perhaps more likely to allow assessment practices, values, and beliefs to expose themselves clearly (Burns & Grove, 1993).

Qualitative research is perhaps analogous to ‘taking a walk through the woods’; I knew where I was starting, but where I ended up would depend on the paths followed. The fundamental tenet of this interpretive approach and the methods that have supported this paradigm is that inquiry is not ‘pre-ordinate’. The researcher does not approach the research with a pre-formulated theoretical stance, but rather begins with individuals and sets out to understand their interpretations of the world around them (Cohen, Manion, & Morrison, 2000). That is why an ‘illuminative’ research strategy was chosen; it was a strategy, which would help to ‘light the path’.

Illuminative evaluation, which is a qualitative and interpretive mode of inquiry, was used as the research methodology in this research. The aim of illuminative evaluation is to inquire into and ultimately to produce an adjudication of a phenomenon by using "description and interpretation rather than measurement and prediction" (Parlett & Hamilton, 1976, p. 88). Illuminative evaluation, being rooted in anthropology (Parlett & Hamilton, 1976), describes and interprets the context in

which educational practices take place. Researchers attempt to adopt an 'insider's' perspective of the individual and explore the activities and discourse of the individual in such a way that they are ultimately able to produce a rich description of the situation under investigation (Fetterman, 1989). Parlett and Hamilton (1976) pioneered this method as a different evaluation approach from the dominant classical or 'agricultural-botany' paradigm of the 1960s.

Burden (2008) suggests Parlett & Hamilton (1972) were reluctant to provide any standard format in which illuminative evaluations should occur. They claimed this process must come from within the diverse forms and differences of educational programmes and the learning milieux. Illuminative evaluation is therefore not a standard methodological package, but a general research strategy which aims to be adaptive and eclectic. The choice of research tactics follows not from research doctrine, but from decisions on a case-by-case basis to decide the best and most relevant techniques; the problem defines the methods used. Equally, no one method is used exclusively or in isolation; different techniques are used to throw light on the problem. The task is to unravel the problem, isolate significant features, and clarify relationships between assumptions and practice.

During the research process the phenomenon to be evaluated is not approached with predetermined criteria for what constitutes 'good' practice. Rather, it is investigated and described as it is, and issues are allowed to emerge as the inquiry proceeds. Issues that are uncovered as significant are then pursued. The basic idea is for the researcher to spend time with the participants (students, and teachers) to pick up how they think and feel about the situation, and what the important underlying issues are (Draper, 1997). In essence, there are three critical overlapping stages: observation, inquiry, and explanation. This approach is based on a number of key assumptions which Burden (2008) explains are:

- Anti-positivist in orientation;
- Thoroughly context bound;
- Multi-faceted in their perspectives;
- Illustrative of the mismatch that often occurs between rhetoric and action;
- Concerned with revealing a recognised and recognisable reality.

This approach involves using a collaborative and participatory technique, which allows the researcher to draw teachers into the process and to drive it. The data is then analysed to reveal 'ideas'. These ideas are then 'progressively re-focused' to develop themes, which lead to meaningful interpretations of the data and clarify the research question(s).

While illuminative evaluation has been criticised for its subjective and impressionistic nature, Parlett and Hamilton (1976) have suggested the problem is not isolated to this approach and point out bias, prejudice and human error are issues in other research methods. They believe that employing certain strategies such as triangulation, spelling out criteria, and openly presenting evidence for others to comment on can help to avoid gross partiality. As such, illuminative evaluation is not value free; instead, its findings represent different value positions, ideologies, and opinions encountered during the research process, which, when represented in a considered and fair way, provide an understanding of complex realities (Parlett & Hamilton, 1976).

Rationale for use of Illuminative Evaluation

The aim of the research was to illuminate the assessment practices of teachers in outdoor education. This was an open-ended question that did not seek specific outcomes or try to prove anything. Two research methodologies were considered as possible research approaches; grounded theory and illuminative evaluation. Both are methodologically similar in that they are descriptive and interpretive (Pidgeon & Henwood, 2004). These methodologies were considered because outdoor education has limited research and it seemed these approaches offered the greatest opportunity to gather broad research data. However, closer examination of grounded theory revealed 'theory' as a summative end to the research. For example, Strauss and Corbin (1998) view theory "as a set of well-developed concepts related through statements of relationship, which together constitute an integrated framework that can be used to explain or predict phenomena" (pg 15). Illuminative evaluation, however, does not seek a definitive outcome or try to produce a theoretical framework.

Illuminative evaluation was first devised initially to evaluate small-scale educational programmes (Parlett & Hamilton, 1972). Since then, according to Rensburg (2007), it has also been applied in the field of health, particularly nursing. Chambers (1988) suggests that this is no accident: as these fields are characterised by complex, unpredictable and messy situations, and the illuminative approach is particularly suitable for these situations. Furthermore, as noted by Harvey and Gunn (2000), an illuminative approach allows the researcher to report on factors important in a particular context and to identify unexpected factors or outcomes. It is an approach that provides support for helping to unravel the issues, isolate significant features, and comprehend relationships between assumptions and practice.

As such, the illuminative approach appeared to align more closely with the research purpose and offered flexibility through this approach. Its research tactics followed not from research doctrine, but from decisions on a case-by-case basis to decide the best and most relevant techniques. Considering this, illuminative evaluation was chosen as the research methodology.

Illuminative Evaluation Approach

There are three recognised stages in the illuminative evaluation methodology. The researcher probes, observes, inquires further and then seeks to explain. The three stages overlap and functionally inter-relate. The aim is to provide information about the ‘instruction system’ (official documents) and the ‘learning milieu’ (the actual learning environment). To do this involves exploring intensively; in this case, the assessment practices of teachers to reveal their practices in order to understand their assumptions, values and processes. To achieve this, the research gathered data through an online questionnaire, observations and interviews.

Data Collection Instruments

These instruments were chosen, for their ability to ‘open up’ the educational situation to intelligent criticism and appraisal, an illuminative evaluation approach (Hamilton, 1976).

Questionnaire

The benefit of using a questionnaire, according to Patton (2002), is that it is self-administered, which allows those participating to respond freely. This provides a useful first step, as it probes the research area by asking general questions about, for example, why teachers assess, what teachers assess, when teachers assess, and how teachers assess. This provides an opportunity for responses to identify and open an area/issue perhaps not considered before. Furthermore, questionnaires are a useful data-gathering instrument for obtaining data from a number of respondents in many locations and are inexpensive to set up. There are however, limitations: participants may not respond to the questionnaire, resulting in a low response rate, or participants may have difficulty understanding or answering the questions. This could provide mixed or incomplete data. However, even with such limitations, this enables the research to begin through a broad-base of data. From this point, the research can then become more focused to gain greater insights in areas that at first may have been unconsidered.

Observation

Observations provide an opportunity to see the participants in real world contexts which offers data that is based on actions, interactions, and judgements. This is essential to uncover actual practices. The analysis of written assessment documentation for example, would be inadequate for this purpose. It also allows another opportunity to gather data by which practices can be understood and to validate the data already obtained from the questionnaire.

Interview

The main benefit of interviews is that they provide an opportunity to gain in-depth understanding, in this case of teachers' assessment practices, and assist in unravelling emerging issues. They provide an opportunity for teachers to tell their assessment stories and for direct questions to be asked.

The three research instruments described follow an illuminative approach, where each instrument builds from the other; this is what Parlett and Hamilton (1972) call 'progressive focusing'. Data is able to be compared against what is already gathered

and this in turn can direct the research into areas unconsidered before allowing problem areas to be explored.

Data Collection

The following is an outline of how the participants were invited to the research and how the data was gathered.

Making contact with teachers

Initial identification of outdoor education tertiary teachers was through a web-based search for New Zealand outdoor education programmes – via http://www.careers.govt.nz/. This identified fifteen tertiary providers. Each tertiary provider's qualifications were then examined through their respective web sites to ensure their programmes and qualifications focused on outdoor education. Five institutes had specialist programmes that focused predominantly on sport or diving qualifications which are not generally considered outdoor education, and were removed from the research. This resulted in a list of ten possible institutes where teachers might be found for the study.

Initial contact was then made by phone to each institute's receptionist to assemble an email list of contact teachers. The identified contact teacher was then sent an email outlining the research and asking for other teachers' email addresses within their institute. Nine contact teachers responded and a list of potential participants was collated. All of these teachers were then sent an email inviting them to voluntarily participate in the research. The email, which contained a link to an online anonymous questionnaire, introduced the research. In total 33 tertiary teachers were contacted. This created a purposive sample group of participants who were experts in their field of practice and able to talk about assessment in outdoor education, telling their assessment stories and revealing their assessment practices.

Table 4.1 Participants involved at each stage.

Questionnaire	15 responses: Anonymous and unidentifiable	**Note:** Some teachers emailed to say that although they were unable to complete the questionnaire, they were still interested in participating in the research. I was not able to observe all teachers interviewed due to time constraints. For the same reason I was also unable to interview all teachers observed. This accounts for the variation of teacher numbers at various stages. In total, 22 teachers were observed and/or interviewed.
Observations	11 teachers: 4 females and 7 males	
Interviews	17 teachers: 4 females and 13 males	

Questionnaire

As discussed the questionnaire aimed to capture a broad and general perspective on teachers' assessment practices. As Parlett and Hamilton (1972) say, the research cannot be charted in advance. Therefore, the main task was to take a complex situation and identify the major aspects before using other methods to explore these identified areas in greater depth. The questionnaire provided teachers with an opportunity to express their initial thoughts about assessment. Capturing this data allowed the research to then develop from the teachers' perspectives.

The questionnaire was trialled before use by a retired outdoor education teacher. As a result of this, a few amendments were made. There were 13 questions (see appendix 1). The questionnaire predominately asked open-ended questions. It was available for completion over a 10-day period to gather initial data from teachers; fifteen teachers responded. The responses were then downloaded and analysed before proceeding to the observation and interview data gathering stages.

Observation

Observation and interview appointments were made through emails. The process of observing and interviewing took place during a four month period and observations and interviews of individual teachers usually occurred on the same day. Some observations were conducted before the interviews and some after. The observations were unstructured, except for the use of a basic guide that was developed from the questionnaire findings. Evidence was sought to further investigate the questionnaire

findings. This unstructured approach was taken because it was unclear exactly what might be observed. The assessment environment was left to unravel teacher practices as the events took place so teacher and student actions, interactions, and judgements were observed as they occurred. The observations were made during practical technical skills assessments. These were assessments of pursuits like kayaking and rock climbing. During each observation, notes were recorded in a field notebook and at the end of each day the observation notes were written into memos. The writing of these memos took place while at various locations throughout New Zealand. This process produced an initial analysis and resulted in a small number of puzzling questions, which were followed up in interviews where possible. Further in-depth analysis took place at a later stage.

Interview

Seventeen teachers were interviewed throughout New Zealand, using a semi-structured approach (see appendix 2). The questions were developed from the questionnaire findings with support from the observations (where these occurred prior to the interviews); as observations occurred their initial analysis added to the interview questions. Teachers were drawn into the research process and as the various stages progressed, they offered increasingly deeper insights. All interviews were digitally recorded and transcribed professionally. The interviews ranged from 45 minutes to 1.5 hours resulting in an average of 12,500 words, which was equivalent to an average of one hour per interview. Transcriptions were reviewed for accuracy and amended as required.

Analysing the data

Questionnaire

The responses were generally short, one-word responses or short sentences. The responses were read and re-read. As key ideas were identified they were noted on poster size paper and clustered together to produce emerging themes. Emerging themes were retained to inform the subsequent stages of data collection.

Observations

Examination of the memos followed a similar process to that employed for examining the questionnaire data. The memos were read and re-read. Key ideas and words were identified and the most frequent ideas were organised together. These were then cross-checked with the questionnaire results before being further focused. The final themes were confirmed once they were cross-checked with the interview data and were reported with the interview findings.

Interviews

The interviews generated considerable data. Because of this, the use of NVivo was considered as a tool for keeping track of various themes, and quotes. However, it became obvious that I first needed to understand the emerging ideas from the interviews in greater detail before I could generate themes by which ideas could be clustered. NVivo was discontinued. Instead, I again read and re-read the transcripts to identify the key and consistent ideas, noting these on each page of each transcript. Each transcript was then reviewed several times and compared with the others, before recurring ideas were reflected on as a whole and placed into meaningful patterns. To manage this process, key ideas were noted on several large poster size pieces of paper, mounted on the wall with support notes and links to quotations (see appendix 3 for examples). These ideas were then cross-checked with the findings from the questionnaire and observation data and only the most consistent ideas were retained.

Writing up the research

The research was then written into draft chapters for critical comments from peers and supervisors and has resulted in this thesis. In the findings chapter the questionnaire results are reported first, followed by three themes, which represent the findings merged from the observation and interview data.

Limitations

The research focused solely on tertiary under-graduate outdoor education programmes (programmes that offer certificates, diplomas and degrees in outdoor education) and because of this, eight limitations presented themselves:

1. Because of the focus solely on the tertiary sector, the findings cannot be considered a cross-sectional view of outdoor education teacher assessment practices throughout New Zealand. They offer insight into tertiary teacher practices only.

2. The sample size was small. The teachers who participated were those who received an email, were motivated to reply to the email, and made time for the research. The teachers who did not respond or were unable or unwilling to participate may hold other views that are missing from the research. The research only sheds light on some tertiary outdoor education teachers' practices.

3. Some teachers may not have been identified and have therefore been unable to participate in the research. Their voices remain unheard, and their practices unobserved.

4. The assessments that I was invited to observe were mainly practical and in the interviews the participants tended to discuss examples of practice assessments. Therefore, assessment practices around theoretical aspects remain largely hidden.

5. What was observed and recorded was inevitably coloured by what the observer pays attention to and considers important, which means the observer may leave unexplored areas that merit consideration.

6. The research is a snapshot of teacher practices. The findings represent a series of short opportunities to capture data and so what is captured could be considered limited. A series of observations and interviews may provide greater evidence and clarity around teachers' assessment practices and result in a richer set of findings.

7. Teachers may have altered their practices just because I was present. This means the data collected may not be a consistent or accurate picture of these teachers' assessment practices.

8. My reputation as an outdoor education teacher with a specialty in kayaking may have influenced what these teachers invited me to observe and what they discussed, hence the focus on practical assessments.

9. Only I have analysed the data, so the findings are my interpretation. It is possible that other interpretations could be made if another researcher analysed the data.

Ethical Considerations

Any research will have ethical dilemmas, and this study is no different. However, not all ethical dilemmas are of equal weight and gravity. Therefore a process to establish the level of risk was necessary.

Firstly, a proposal documenting the intended research purpose, rationale, and significance was completed. This included details about how the research would occur, who the participants would be, and discussion around the identified ethical issues. The main identified issues were:

- the need to minimise risk of harm to participants;
- the need to provide informed and voluntary consent (appendix 4 & 5);
- the need to be respectful of privacy and confidentiality;
- the need to avoid unnecessary deception;
- the avoidance of conflict of interest.

To establish the level of risk, the Massey University research ethics screening questionnaire was completed and this helped to define the research as low risk. Together with my research supervisors, we further reviewed the risk-screening questionnaire and agreed that this research was low risk and should be registered with Massey University Human Ethics Committee as such.

Management of the key identified issues

Minimising risk of harm to participants

The risk of harm was considered as a potential issue, as I would be involved in making a number of direct observations where teachers were assessing students. It could have affected the student(s) being assessed through:

a) my presence being perceived as an added indirect pressure;
b) a lack of clarity around my role;
c) interference with the process through comments or inappropriate behaviour.

To mitigate these possible intrusions, I discussed my role with each teacher and asked them to clarify the situation with their students. They were told that I was only present to act as an observer of the educational assessment process and that I held no authority and had no intention of making judgements about any student's skill, knowledge, and capabilities. My role was to remain as much as possible a bystander, out of the way. My behaviour was consistent with normal expected teacher behaviour and followed good teaching / research practice guidelines (Massey University, 2009).

Respect for privacy and confidentiality

This was considered essential to prevent the identification of any teacher. All tapes, computer discs, and paper copies of data were kept in a secure place. This was managed by downloading all data on to my personal computer with password protection. Participants were informed that all electronic data would be erased once the research was completed. Access to transcriptions of audio-taped interviews would be available to the researcher or a trusted person for transcription purposes (who agrees to keep the information confidential) and the research supervisors. To maintain confidentiality, any discussions about the research avoided specifics to protect the identity of individual teachers, which includes the institutes where the teachers were employed.

Avoiding unnecessary deception

All participants were provided with an information sheet which informed the participants what the research was about, its intended use, and what involvement was required. Before any observation or interview took place participants were asked if they had any further questions about the research and reminded that they could

withdraw from the study at any time, or decline to answer particular questions. Participants were then asked to sign a consent form acknowledging they understood the research and their rights. They were informed that the research would be used only for this master's degree and publications arising from the research study.

The avoidance of conflict of interest

The key identified conflict was the involvement of work colleagues. To prevent this all CPIT colleagues were excluded from the research. In essence, the research was about understanding professional activity and individuals had the right to decline to be involved in this research.

After these ethical issues had been discussed with my supervisors and strategies put in place as outlined above, the research was registered with Massey University research ethics committee as low risk, and accepted as such (see appendix 6). Coincidently, during 2009 I was granted academic study leave from Christchurch Polytechnic (CPIT) for the start of this research. CPIT also required the research be registered and screened for ethics approval. The research proposal was also approved by the research committee of CPIT.

Chapter Five: Research Findings

This chapter extrapolates the data from the three stages of this research. Firstly, the initial data gathered from the questionnaire is documented. The chapter then identifies three themes, the most frequent identified from the observation memos and interview transcripts. The first theme, *Perceptions of Assessment* illustrates some of the challenges and difficulties teachers face around their assessment practices, while highlighting the value teachers place on learning. The second theme, *Assessment Focus*, illustrates the concentration teachers have around technical skill assessment. This emphasises also identifies neglected areas of assessment in outdoor education practice. The third theme, *Assessment Criteria and Professional Judgement in Assessment*, highlights teachers' perceptions of criteria as standards for assessment and the use of professional judgement as an important aspect in assessment decision making.

Questionnaire findings

There were 13 questions, which were predominately open-ended questions, where participants could describe their interpretation of the questions asked as it applied to them and their context. Questions 5, 6 & 8 were closed questions where a range of replies were provided, which allowed the participants to rate how often or not this applied or occurred in their practice (see appendix 1 for detail) . Fifteen teachers responded to some of the questions. The most frequent responses are documented below.

Q1. Why do you assess?

Fourteen teachers responded to this question. Three themes emerged:

- Standards monitoring and measurement (7 responses);
- Feedback to students, teachers and third parties (7 responses);
- Legal obligation (7 responses).

Teachers said they assess summatively because they are required to, but also to maintain standards. Formative practices were also frequently commented on for their contribution to learning and improving teaching practices.

Q2. What are the five most important aspects within your Outdoor Education programme that you assess?

Fourteen teachers responded to this question. Five aspects arose as the most important for outdoor education assessment; they appear in descending order according to how frequently they appeared in each teacher's response:

1. Technical skills (11 responses);
2. Attitudes (7 responses);
3. The ability to think (6 responses);
4. Knowledge (5 responses);
5. Judgement (5 responses).

Most teachers placed a strong value on teaching and assessing technical skills in outdoor pursuits. Although they were aware other aspects were essential, they appeared to have a lower value.

Q3. What makes these aspects important?

Thirteen teachers responded to this question. Two key findings arose from these responses:

a) Technical skills were justified because they provide a foundation for safety, although teachers commented that being technically, highly skilled did not always provide the best teachers in the field.
b) Teachers also acknowledged attitudes, judgement, and social skills were foundational for effective teaching and developing people through outdoor education.

Response (a) above is congruent with the first response in question two, however, response (b) above appears contradictory with the other results from question two. Assessing soft skill aspects like attitudes and judgement were given a lower rating in question two than teachers appear in this question to rate their importance.

Q4. How do you assess these important aspects? and

Q7. What methods do you use to assess learning?

Thirteen teachers responded to these questions. These questions were linked together as they produced similar results.

Most teachers used a wide range of assessment strategies: practical assessments, written assignments, observation, group tasks, gathering naturally occurring evidence, examinations, oral questioning. Their practices also included self and peer assessment. Interestingly a few teachers commented that they, “knew who will and will not pass and why, long before the "test" or assessment occurs”.

Q5. When during a course do you carry out assessments?

Fifteen teachers responded to this question.

Over 70% of these teachers used summative assessment at the end of their courses or throughout their courses. This result is puzzling, since all teachers would be required to conduct and report summative assessments. This raised the question what were the other 30% of teachers doing about summative assessment?

Predominantly summative assessments occurred when they were scheduled, and on rare occasions, they occurred when students were ready.

Over 50% of the teachers used formative assessments at the beginning of their courses and this practice is continued throughout the course. Thus the practice of assessment *for* learning was prevalent, whereas, summative assessment appeared to remain a one-off, high-stake end event.

Q6. Does your teaching and assessment include any of the following?

Fifteen teachers responded to this question. Their answers are noted in descending order according to how frequently they appeared in the teacher responses:

- 92% said they matched their assessments to the learning outcomes;
- 86% generally used explicit criteria for assessing;
- Close to 50% pre-moderated their assessments and just over 50% post moderated their assessments;
- Blind and double marking hardly ever occurred;
- The use of professional judgement in assessment was very common practice.

Q8. In your teaching who is involved in making assessment judgements?

Fifteen teachers responded to this question. The stand out response was that over 90% involved students in assessment judgements, most of the time.

Q9. What aspects of assessments work well for you?

Thirteen teachers responded to this question.

This question produced varied responses. Three themes emerged:

1. Involving students in the setting of assessment criteria;
2. The use of authentic assessment tasks / situations;
3. The use of naturally occurring evidence.

Teachers said that they tried to use formative assessment and avoided the use of high stakes summative assessments. This result was puzzling, as summative assessment cannot be avoided if results of a course are required to be reported, and further more this contradicts the findings from question 5, where 70% say they run summative assessments.

Q10. What problems do you experience in assessment?

Fourteen teachers responded to this question.

This also generated a variety of responses. Six themes emerged:

1. The difficulty of assessing attitudes;
2. Deserving students failing for several reasons: minor faults; becoming stressed and underperforming; getting hung up on the final assessment, in spite of knowing the material or being able to perform the skills under formal assessment conditions;
3. Teachers' inconsistency in assessing standards;
4. Managing high volumes of students;
5. Gathering sufficient evidence;
6. The time consuming nature of practical assessments.

The difficulty in assessing attitudes helps to explain the results in question three, where attitudes were said to be valued but held a low rating in question two in terms of importance.

Q11. Other comments?

Seven teachers responded to this question.

There were no consistent themes. However, two comments are worth noting,

“Assessment can drive learning rather than the other way around. Too much focus on credits, ring-fences what the student needs to know and doesn't encourage them to learn outside that area” (participant 2).

“Assessment is a teaching and learning tool, not an end in itself” (participant 7).

Q12. What industry qualifications do you hold if any? and

Q13. Demographic information.

This purposive group of outdoor education teachers were well educated. Over 80% of them held tertiary qualifications. See tables below.

Table 5.1 Participants educational qualifications

Number of teachers	Type of qualification held
3	No formal post secondary qualification
3	Diploma
8	An undergraduate degree
1	A post graduate certificate
1	A Masters degree
1	A PhD

All the teachers also held a number of the highest and most recognised industry awards, namely NZOIA instructor awards, but none indicated they held Skills Active ITO qualifications. Other qualifications such as first aid, avalanche, or heavy vehicle licenses, which are also common outdoor education teacher qualifications but are not specific to outdoor education were not included in the findings. The table below illustrates the participants’ industry qualifications and shows the diversity of skills and experience these teachers bring to tertiary outdoor education training programmes.

Table 5:2 Participants industry qualifications

Industry qualifications	No. of teachers holding each award
Mountain Stage 1 instructor award	5
Mountain Stage 2 instructor award	4
Rock Stage 1 instructor award	11

Rock Stage 2 instructor award	5
WW Kayak Stage 1 instructor award	13
WW Kayak Stage 2 instructor award	6
Sea Kayak stage 1 instructor award	6
Sea Kayak Stage 2 instructor award	3
Bush Stage 1 instructor award	11
Bush Stage 2 instructor award	10
Raft assessor	1
Snowboard stage 2 instructor award	1

Furthermore, half of this group were national assessors for NZOIA, which means they were part of a small group of experts, who set and judge standards for professional outdoor education teachers / instructors throughout New Zealand.

Table 5:3 Participants experience

Average number of years involved in outdoor education delivery.	**Average number of years involved in teaching outdoor education at a tertiary level.**
17 years	7 years
A range of 7 to 39 years experience	A range of 1 to 15 years experience

It can be seen that extensive industry experience underpinned the teaching practices of the participants. However, the group had mixed levels of formal teacher education qualifications and training and varying awareness of teaching, learning and assessment theory, processes, concepts and principles. Seven teachers had completed a diploma in teaching (adult or secondary) with a further three having completed a certificate in adult teaching. This means that just over half of the teachers had some formal professional teaching and learning education.

Observation and Interview Findings

The questionnaire provided an initial probe into what, where, when and how teachers assess outdoor education and the findings above document analysis of the results. These findings were then used to explore in greater depth the assessment practices of outdoor education teachers. The following section is an analysis of observation and interview data presented in the form of the themes identified. Teachers' quotes are placed in *italics* as is the referencing of these. For example, *(6: p12)* refers to teacher 6 and the quote is from page 12 of their transcript.

Theme 1 – Perceptions of Assessment

Perceptions of assessment were varied, ranging from a view where assessment was seen as a difficult process separated from teaching and learning to one where it was seen as an integral part of teaching and learning.

Many teachers perceived summative assessment as an uncomfortable, difficult, and complex process, which was full of challenges.

> *I don't enjoy assessing that much, I mean I do assessments because I have to provide evidence of where the students are, to pass them or not (17: p5).*

> *It is a multi-headed monster (16: p 7).*

Moreover, summative assessment was seen as a requirement, not a helpful process, to aid teaching and learning. Another example illustrates similar responses.

> *I often have frustrated thoughts around assessment as it is overly constrained by things such as achieving 100% competency (5: p4).*

The failure of students to demonstrate competency in an assessment, caused frustration because it was felt that, the benchmark sometimes had to be altered to allow students to pass to maintain pass levels.

> *You need to lower the standard to where you want the pass level to be at, to allow other students who have done bugger all, to pass (5: p4).*

However, observations of assessment practices did not detect this. Yet many teachers made this point. Observation revealed students performing skills at a range of levels, and teachers' subsequent judgements, assessed students as competent or not yet competent. The observed difficulty for teachers appeared to be deciding where the benchmark lay. While many of these teachers appeared to be astute judges of student

performance, when it came to defining the absolute minimum pass standard, they had trouble deciding this benchmark.

Teachers felt they were being pressured from various stakeholders, especially students and institutions, and were aware that assessment was affected by competing influences. One teacher observed:

> *The students complained about the stuff that I was teaching because there was no assessment tagged to it. They wanted stuff to be assessed, saying, "I'm not learning this unless you assess me" (16: p8).*

This was not unusual. Teachers felt that students were only interested in learning if it had direct relevance to the assessments of their course. This forced them to think they had to teach to a narrow set of objectives and that learning for intrinsic values was being lost. The result was many teachers re-wrote assessments to include all aspects they taught. However, this resulted in excessive assessment, which meant many of these teachers simply reduced the breadth of their teaching.

Another perspective was that perception that institutes perhaps valued pass rates as more important than standards or the quality of the teaching.

> *The tertiary provider wants to have a measure of quality in their programmes, and their measure of it is how many students pass the assessment. All papers must therefore have some component of assessment. It doesn't matter what the assessment is, as long as they get a percentage of students passing it (16: p8).*

Although teachers found assessment difficult, several teachers viewed summative assessment as enhancing learning and supported the notion that it can contribute to learning.

> *Assessment can be such a powerful learning experience (10: p19).*

Teachers acknowledged that assessment can be a motivator and that it encouraged students to learn and provided an incentive for them to prepare for an assessment.

> *I do see the flip side of assessment. It is a real motivator for the students if they know they have to be at a certain standard and they are going to be assessed. I am sure some students are driven by that, so it is important to motivate students (17: p5).*

Teachers also said there were a number of positive learning benefits such as:

> *motivation, co-operation between students, planning, and self-directed practice (17: p5).*

Another perspective saw summative assessments being valued for preparing their students for industry assessments, and saw in-house summative assessments as less important.

> *Because we are such firm believers of the NZOIA system, we don't see our assessments having to be super rigorous. Summative assessment tends to get them* (students) *really stressed and bring out the worst in them, which is another reason why I would try to de-emphasise it (7: p10).*

Observation evidence supported this view, and many teachers created a relaxed assessment atmosphere to help de-emphasis summative assessment. This allowed teachers to provide feedback during summative assessment, so they could continue assessment *for* learning at the same time. Thus, student learning remained a central priority during summative assessments. There was a real sense that all of these teachers cared about their students, and the observations demonstrated formative assessment practices were a natural routine.

> *I certainly see formative assessment as an important part of learning (7: p6).*

> *I am always driving people to get better, do more, try more, be more motivated....I just want them to spend more time getting better and that is constantly in the back of my mind. So, yes, I don't like assessing particularly. I'm just wanting to teach people (6: p8).*

Teachers clearly value the role of formative assessment and give learning a high priority.

This raised an important question for some teachers where they questioned the role of summative assessments in tertiary education.

> *There is a tension between tertiary education and the continuation of a lifelong learning journey. Tertiary assessment is only a point in time and that has a half life (3: p5).*

Teachers thought tertiary education created a culture where assessment was seen as an end to their learning.

Assessment is about students getting their qualification and culturally that is how it's been put forward (3: p5).

However, many of these teachers were promoting a culture that encouraged ongoing learning,

I'm trying to give the student feedback, trying to help them with their journey. I believe that they're still on a learning curve, even during their assessments. And then after their assessments they carry on with professional development (3: p7).

This perspective highlights the tension some teachers perceive between assessment for a qualification and the ideal of a lifelong journey of learning and professional development. The later ideal led teachers to consider new ways of practising assessment, one where students play a role in the learning and assessment process.

I aspire to greater involvement of the students in the process of their own learning. I recognise that there is some good stuff around that and would like to increasingly involve the learner in the process of shaping their own learning and, by extension, their assessment as well. But at the moment there is not a lot of this. However, I am philosophically aligned to it (12: p26).

This illustrates a change in the way some teachers perceived their own role and that of their students role in teaching, learning, and assessment. Some teachers are already trying to find new ways to incorporate greater student involvement in their programmes, for example,

*Students assess themselves. There's a facilitated debrief session. They created their own criteria, and set their own benchmark. I've given them guidelines about our standards and they created their own criteria. I ask them how well they performed as a guide. They score themselves out of five...*and you run with their judgement on that? *Most of the time I do (14: p17). It's actually quite refreshing for the students to be able to do it* (self assessment) *for themselves. I'd do this more if I had time (14: p19).*

It is clear that teachers perceived student involvement aided the learning and assessment process in a positive direction, developing a culture of self-responsibility and self-learning.

Theme 2 – Assessment Focus

Observations of and interviews with many of these teachers revealed that they focus primarily on assessing technical skills, outdoor recreation pursuit skills. These skills, were predominantly about performance and rescue capability in outdoor recreation pursuits.

> *Our assessment in the certificate course is basically about personal technical skills (13: p10).*

> *There has been more of an activity focus because the activity skills are the ones that form the bulk of the kind assessed and there is still a heavy focus on activity delivery in the diploma programme (12: p24).*

All the observations involved the assessment of technical skills in outdoor pursuits. At no stage during observations were students assessed delivering outdoor education with clients. There were no examples of assessment of a student's ability to teach outdoor education and apply soft skills and other aspects such as judgement, facilitation skills, and decision-making.

However, after initial discussions, and with probing questions, many of the teachers were acutely aware that there was more to assess than just technical skills.

> *Student judgement, the attributes of a guide, making good decisions, are some of those other sorts of competencies. As are the likes of communication skills, the soft skill side of outdoor education (17: p8).*

Teachers described concern about and awareness of other aspects that should be assessed. They said that these accepts were less tangible than technical skills, but were equally important, acknowledging that they were more difficult to assess. Teachers hoped their students would acquire all the professional skills needed to offer safe and effective outdoor education experiences. As such teachers asked themselves question like the one below:

> *Would I be happy to give you my 10 year old daughter to go and do this activity with this student in charge? Could they do what they are doing safely (2: p8).*

This broad and open-ended question may be an important question, yet there was no formal summative assessment that assessed other aspects that make up safe and effective outdoor education.

As soon as you are talking about delivery and instruction of outdoor education, then you implicitly know you are talking about more than activity, more than technical skill. Yet they are reasonably vague performance criteria on those, as they probably inevitably will be really difficult to assess. Because leadership, for example, is still a reasonably nebulous kind of concept in many fields (12: p24).

Many of these teachers suggested that other aspects, for example, leadership, were assessed when students delivered outdoor education to clients.

I feel as though I am assessing soft skills but I would struggle to show you the criteria that I am assessing it against (p25).

However, none of the observations saw students being assessed in this situation. From the interviews, it appeared that little assessment occurs where students were teaching outdoor education to clients.

While many of the teachers could describe what they meant by the less tangible aspects of good outdoor education, they found it difficult to document and provide criteria. The following comments reflect a common solution.

The less tangible aspects certainly appear in our reference at the end of the year. The character personality stuff I think is really important and it is also difficult to assess. One of the things I am quite keen on is that the references we write at the end of the year are about the real person rather than someone that they are trying to be. For example, in kayaking, there are some holistic tick boxes like 'is a responsible group member'. But it is probably more key events....I was particularly pleased today to see a bunch of boys all get out of their kayaks and go with their throw bags and support the next group coming down. So yes, that is a little tick which will get remembered. If any of them at the put in today didn't have their throw bags that is a little cross that would certainly get remembered (7: p 17).

Although teachers tended to perceive the assessment of the less tangible aspects, as difficult, overly subjective and not able to be defined well enough to enable a fair assessment, in many cases however, they expressed a desire to assess these other aspects. In practice, however, explicit learning outcomes did not reflect this expressed desire.

Assessment is more than a kayak assessment. The criteria don't really help you, because the criteria are all based around: Can you kayak? or Can you climb? or

Can you do this and the next thing? And that is reasonably effective in assessing that they have got those skills. We are confident that we can turn out people who are skilled but there is all that other stuff that is in the curriculum document. You are constantly assessed on your professional attitude and your attitude to learning and da de da (6: p10).

Moreover, teachers' individual courses, which made up their programmes and the learning outcomes and subsequent criteria for assessment, did not align with the graduate profile. The syllabuses simply failed to note the other less tangible aspects, instead focusing on technical skill competence.

There is nothing in the unit or assessment schedule that talks about soft skills specifically (6: p10).

Most teachers perceived the assessment of pursuit technical skills as measurable, compared to other aspects. Yet, in stark contrast, one teacher working with a number of other teachers within their programme, clearly defined some of these aspects.

In all our programmes there is an attitudinal outcome so they need to be on their game. We tell them this is a workplace and in relation to attitude it is about turning up on time, being clean and tidy, being nicely well presented. We have grading criteria which are partly based on attitude, and there is list of things that we look for as far as attitude is concerned (14: p9).

Being tidy and clean may be a limited view of attitudinal outcomes, however, it illustrates that it is possible to identify and define less tangible aspects of outdoor education for learning and assessment purposes. If these aspects could be more clearly identified and written into learning outcomes, then assessment criteria could be developed to support learning and assessment in outdoor education. This does not mean it is an easy task, as was highlighted in the questionnaire results.

It's a tough game (assessing attitudes) *it has a lot of subjectivity and there's a lot of kind of professional judgement involved (15: p16).*

Even if criteria were presented for teachers, the use of professional judgement was seen as subjective and raised issues around assessment reliability and validity.

Theme 3 – Assessment Criteria and Professional Judgement in Assessment

Two possibly contradictory assessment practices presented themselves in this theme. One was the use of stated criteria to enable students to understand the required assessment standard and to ensure assessment was as objective as possible. The other was the use of professional judgement to assess student work or performance to enable an assessment decision to be made. The latter practice was, however, considered less transparent, and perceived as overly subjective by many teachers.

Many teachers believed their assessment criteria indicate the standard for each assessment, and referred to these standards as a benchmark.

> *There is a list of skills that they get assessed on and then there is also a description of essentially what gets looked at (7: p21).*

However, closer examination of criteria sheets during field observations revealed that the criteria were not that well defined and were, instead, vague and quite open-ended. For example, teachers said that defining criteria was:

> *pretty difficult really. You can put it* (criteria) *into words but that doesn't tell you a lot. For example, I use words like 'smooth', 'effortless', 'efficient' - those sorts of things I guess I'm looking at. But they still require interpretation (13: p12).*

This illustrated the problem, as the criteria do not state the required standard clearly. The criteria needed further clarification to be of any use. Several teachers recognised this issue, especially when setting criteria for fluid and dynamic technical skills assessments.

> *I have always struggled a bit with the whole idea of assessing these things technical skills to a tick box list. The process is hard to do unless you have an objective degree of skill measurement (6: p6).*

Defining objective criteria for technical skills in outdoor education is not a quantitative process as is found in sport science. This is because outdoor education performance is not about improvement for competition. Criteria are by nature qualitative as the assessment focuses on students being able to produce images other could learn from in their future roles. Simply, outdoor education does not focus on judging skills for measuring speed or strength. This clearly causes teachers challenges as they set about trying to define criteria for standards.

Teachers commented that the development of their assessment criteria occurred somewhat haphazardly. A few teachers said they simply inherited pre-existing criteria.

> *I guess a lot of it had been set before I arrived. They were already existing (1: p11).*

> *I'm buggered if I know how they were developed. I haven't had time to look at them, literally (5: p10).*

> *I made it up as I have gone along, you know. I truly had no real idea of what I was doing. The Certificate in Adult Teaching course has helped. But for me it has been a process of looking at what has gone before and thinking, well, that doesn't really make sense, and just changing it. I have only just figured out what Bloom's taxonomy is in the last six months. So I had no idea (6: p25).*

Some teachers appear to be lacking in good assessment knowledge and how to go about establishing valid assessment criteria. Others were clearly 'deep-ended' into an educational assessment with little support or mentoring. Most teachers said that their criteria were developed in isolation, as this example highlights:

> *Often we work as little autonomous islands - we're doing our own thing and sometimes that's good and sometimes that's bad. The development of criteria is just a personal, subjective thing. Sometimes we have conversations, but usually it's done in isolation (16: p29).*

These examples reflect a possible lack of teacher training. It appears some criteria were developed in a void of support or consultation with other teachers and with little learning and assessment theoretical understanding. The result being that the criteria were unrefined and one teachers subjective view of the expected standard. This is perhaps amplified because according to some teachers:

> *There is no real standard out there that we can work to, there's no other benchmark out there. So now I try to employ staff who are working off their NZOIA qualifications (14: p16).*

Some teachers may hold NZOIA qualifications in high status. Yet, ironically, several teachers recognised that

> *the NZOIA benchmark standard is quite hard to communicate because it isn't written down. They just give you a syllabus (13: p12).*

Therefore, while many teachers were partly employed for their NZOIA qualifications, they were trusted as professionals to set standards based on their qualifications, experience, and judgement.

Given that the criteria were vague, most of these teachers did not commonly assess using checklists. Teachers said checklists pulled everything apart, where things like kayaking and climbing required a number of skills and techniques to be linked together to create a whole. Instead, criteria were a general description of what a student should be able to do. The actual criteria however, seem to be held in teachers' heads. They assessed through their professional judgement, believing they were astute judges of standards.

> *You just have a picture of what the person's performance should be like, I hold a picture with me. I know what they should look like, as I have seen thousands of these now, and I am an astute judge of when they are working well (6: p31).*

The significance of this for students probably means they have minimal specific criteria, which makes it difficult to understand the standards expected. Regardless, all teachers made an effort to ensure students understood the assessment standard and expectations.

> *I think it is really important that the students know what they are being assessed on probably even before they step up for the day. They are given the criteria, you know, on the first day. You can't expect people to guess, I am into being fair (4: p7 & 8).*

Teachers also dedicated considerable time to helping students understand the assessment criteria.

> *I'm not into giving people a marking criteria and a course outline and leaving it at that. We give our students course outlines and with it the marking criteria. And so I'm not expecting them, without any more input from me, that they will understand it. So I believe that there is a continued process (3: p13).*

It is, however, not just having the criteria that are important.

The person that's being assessed needs to understand what's expected of them. They have to have some understanding of the criteria of what you're looking at and why. And that's quite a challenge really, because you give someone a bunch of words on a piece of paper and what they interpret from those words can be completely different from what you expected those words to mean to that person (3: p10).

The challenge is how to communicate criteria to students to provide them with a fair opportunity of meeting the intended standard. Some teachers were well aware of this issue and managed this by:

re-addressing the marking schedules and go through it and talk about my interpretation from it, to help them understand what I wrote. It gives them a chance to question and say "So you mean this?"(3: p10).

Many of the teachers argued that students learn about the assessment standard through their direct involvement with the learning programme and through direct discussion about the criteria. There was a balance sought between explaining criteria and understanding what is expected through the teaching and learning programme.

Students have been out and they've seen instructors instructing. So hopefully that means they've seen some good instructing and they know that roughly this is the kind of thing they're hoping to do, a big picture kind of experience (10: p28).

It is not just a matter of just reading it in the course outline. I guess they also get understanding during the learning process and things like good demonstrations, so they know what they're meant to look like themselves (1: p10).

I think the students get more of an indication of what they will need to know from what they are taught rather than reading the criteria because criteria don't have the standards in them (5: p18).

It is clear therefore, that teachers believed students learnt the standard expected for assessment through a range of different encounters with the criteria. From observation and interview discussions, what became apparent was that the was agreement on standards among teachers, even when standards were mostly written in isolation from other teachers.

I have got so much experience working for NZOIA and alongside other people in a professional capacity that you have kind of built up a picture of or a vocabulary of what things should look like based on that experience (7: p15).

The more our staff can actually talk together, the more moderation and dialogue we have comparing personal judgements and the more we start talking about the 25 pages of criteria in some sense (3: p18).

Teachers talking and comparing standards agree through their professional experience. This is perhaps what

drives not writing so much criteria. I think it is the economics of effort and time it would take to actually write a lot of the criteria, you know, no one is saying to us you need specific criteria, there is no pressure to have those criteria in hard terms (3: p29).

I've got screeds of paper from Skills Active that lay out what 'competent' looks like and what kind of evidence supports that. In fairness I do not spend a lot of time pouring over those. I read them, get a sense of it. A good teaching day of rock climbing, gives me, I think, a much clearer picture of the competence of the second year students, than if I set up a bunch of assessment tasks with set criteria (12: p16).

From this it is clear that professional judgement is an essential assessment tool that is widely used. There seems to be a feeling that written criteria tend to hinder a more holistic view of assessment and professional judgement works in conjunction with criteria. Ultimately, most teachers appeared to have faith that their peers are doing the best job that they can based on their professional background.

These findings in summary appear to highlight four key issues:

1. Many teachers found summative assessment a difficult proposition, especially as they delight in teaching and learning.
2. The assessment of technical skills was well established; however, assessment of other aspects of outdoor education were limited, even though they were said to be of importance.

3. Criteria for assessment were somewhat vague, and often lacked the clarity they should offer a student trying to understand the assessment standard.
4. The use of professional judgement to inform an assessment decision was considered overly subjective and not used with confidence.

The next chapter discusses these issues, suggesting possible reason why these issues are present, and explores potential solutions with literature support.

Chapter Six: Discussion

In this chapter the findings are considered to illuminate their meaning, discussing the practices, values and assumptions teachers may hold in outdoor education at a tertiary level.

Teacher Assessment Competence

This purposive group of teachers are highly educated, and hold numerous NZOIA industry qualifications. They are a select group of expert outdoor education teachers, and many of them were national industry assessors with extensive experience. The expectation might be that such teachers would also be highly trained in teaching, learning and assessment practices. This was not necessarily the case; teachers sometimes did not have additional adult education qualifications, as is often the case in tertiary education, and this raised questions about their assessment literacy. Many had learnt about assessment through trial and error and because of this there may be large gaps in their comprehension of assessment theory, terminology, concepts, and practices. This might also help to explain some of their struggles with summative assessment and the undervaluing of formative assessment judgements for summative decisions.

Assessment literacy, as Stiggins (2002), and Volante and Fazio (2007) say, involves the understanding and appropriate use of assessment practices along with the knowledge of the theoretical and philosophical underpinnings in the assessment of students' learning. DeLuca and Klinger (2010) suggest that because current assessment practices emphasise criterion-referenced and standards-based assessment, there is an increased need for teacher competency in assessment. Teachers must be able to justify their assessment decisions and be accountable for them. Teachers who lack in-depth understanding of assessment may be unable to engage in meaningful and critical debate around the topic. DeLuca and Klinger's (2010) suggest, otherwise this may lead to poor assessment practices where the validity and reliability of the assessment is reduced, resulting in inappropriate educational decisions. The most probable outcome of this is the replication of teachers personal past experiences of assessment. This is most likely to be a summative assessment model that only measures learning and grades students. However, for these teachers, this is not their

default position, they clearly use formative assessment techniques. Nevertheless there does appears to be some opportunity for professional development in assessment theory and achieving summative judgements through varied and accumulative methods instead of relying on one-off assessment events.

Teachers Perceptions and Practices of Assessment

There is no denying that summative assessment in most cases is a complex process, full of traps and issues. Even the use of criterion-referenced assessments can cause frustration because students who have performed with greater skill and know more, are not identified and receive the same result and end qualification as their fellow students. However, as Biggs (2003) explains the beauty of criteria-referenced assessment is obvious: "Say what you want students to be able to do, teach them to do it, and then see if they can in fact do it" (p. 144), however, these teachers would contest that assessment described in this way is as straight forward as Biggs proposes.

Given the difficulty and frustration teachers can experience in summative assessment, it was surprising to observe that most teachers did not use checklists because lists with the components of competence would seem a useful tool for assessing a range of technical skills. However, teachers felt their use created a culture that reduced complex aspects of technical skill performance into small isolated elements and that, '*absolutely killed the experience and lead to a passionless, uninspiring day' (6: p20).* Biggs (2003) and Sadler (2007), both support the practice of not using checklists. They argue that separating components of performance into isolated elements, rather than looking at the whole brings into question the validity of the assessment.

Assessing holistically allows technical skills to occur naturally as a normal experience, gathering evidence of competence as the assessment unfolds, a process that is referred to as naturally occurring evidence. However, the absence of checklists may cause other issues; for example, some aspects of competence could remain un-assessed if the criteria are not written down and no schedule followed. Therefore,

both reductionist[5] and holistic[6] approaches to assessment present issues for validity and reliability. Whatever the model of assessment used, all approaches to assessment are likely to have limitations for teachers, students, and stakeholders. The best assessment method according to Biggs (2003), is one that best realises the course objectives, which should also align to the graduate profile.

A key aspect of this may be the position summative assessment has in the teaching and learning programme. Summative assessments appear to occur predominantly at the end of a course of study. This creates a situation where, as Hume and Coll (2009) pointed out, assessment becomes a one-off, high-stakes event: the dreadful final assessment or examination. It requires students to execute skills and knowledge on demand and perhaps in artificial situations. In some situations this may cause them to underperform. Such practice denies formative assessment observations and ignores naturally occurring evidence that could be used when students were engaged in prior practice and learning. This naturally occurring evidence can be a valuable support for teachers arriving at valid and justifiable summative assessment decisions. Teachers openly commented that they had often seen their students perform the necessary skills many times beforehand; however, these observations as legitimate judgements for summative assessment purposes were largely ignored.

Light et al. (2009) point out that one barrier to change in the area of assessment is the way teachers understand the role and meaning of assessment. According to Watkins et al., (2005), one of three points of view about the role and meaning of assessment is often prevalent. The least sophisticated understanding of assessment is where it is seen as a separate process, external to the teaching and learning. The second conception of assessment is an awareness of the internal relationship between assessment and teaching. This perspective views assessment as essentially being about measuring students' acquisition of basic knowledge. The most sophisticated conception views assessment as an internal and integral aspect of teaching and

[5] **Reductionist or reductionism**
Is an approach to learning and assessment where complex skills and knowledge are reduced to a set of simpler components where the parts are separated from each other.

[6] **Holistic or holism**
Is the idea that everything exists in a relationship, in a context of connection and meaning. To separate them in learning and assessment loses these relationships.

learning, which focuses on developing deeper levels of learning, where teaching and learning involve practice which promotes understanding, critical reflection, analysis, interpretation, and applying knowledge to solve problems. These teachers appeared to employ the second conception of assessment, focusing on measuring the acquisition of technical skills. Given teachers have often seen students perform technical skills through formative phases, summative assessment could be integrated into their teaching and learning programme with minimal adaption. Formative assessment judgements could, therefore, help to provide judgements for summative decisions.

The support for final assessments was perhaps an unconscious one in that teachers may have unthinkingly continued assessing in the same format as they themselves were. It can be argued that summative assessments where students are placed under stress may be a useful tool to establish whether they can manage stressful conditions. This may well be a desirable outcome, given the nature of their future work, where risk is involved and where life and death decisions may need to be made. Yet there was no evidence that the ability to cope with stress was an assessment criterion. If the goal of summative assessment is to see whether a student has achieved the outcomes of a course, then the assessment should be aligned to the learning outcomes and should assess those outcomes through authentic tasks. If the ability to work under stress is to be assessed, students should be aware of this and it must be stated explicitly. A one-off summative assessment of this kind is also unlikely to be a valid or reliable assessment, as it is doubtful that it would provide sufficient evidence. Furthermore, one-off final assessments may mean a student flukes the assessment, raising questions as to whether a student could reproduce the same result another time or in variable situations or conditions. This introduces an element of chance into assessment. Teachers almost certainly need to consider if or how their assessments are affected by a range of factors: for example, what effect the weather had on the students' performance or which aspects were assessed and which ones were not. Some teachers were aware of this issue and used judgement in an attempt to make summative assessments between different assessment conditions as fair and consistent as possible.

Assessments basically involve some degree of risk, and this requires teachers to balance the act of gathering sufficient evidence and evaluating this evidence fairly and consistently. However, as Crooks (1988) says, it is important to avoid the error of substituting more reliable measures over validity measures, as validity is clearly more important. Sadler (2007) suggests assessment may sometimes benefit from accumulative or progressive assessment judgements and that formative assessment can be used during the learning phrase for summative purposes. One-off final assessments that focus on a multitude of discrete competencies, he argued, should be avoided. Biggs (2003), also argues that separating out parts of practice independently and then assimilating them does not provide an assessment which can be considered authentic. A judgement of competence should be considered an overall judgement of a student's capability: skills, knowledge, attitudes, judgements, and decisions in real-time situations. A change in the way teachers think about assessment and gather evidence could provide a broader range of evidence, which may well offer better support for more reliable assessment judgements.

While the immediate observations and interview discussions saw teachers focus on summative assessment, their formative practices - assessment *for* learning were highly visible. It became apparent teachers' normal practices were formative, and promoted a culture of learning. This helps to explain further the difficulty teachers experienced with summative assessment: they would rather be teaching and helping students learn than conducting summative assessments. It was clear that teaching and learning were highly valued and summative assessment often was viewed as either failing to assist learning, or getting in the way of learning. It should be noted that some teachers, however, were aware that summative assessment could assist student motivation. They commented that students could all of a sudden become very motivated to learn, practise skills, and improve their capability, due to the demands of impending summative assessments. So, as much as summative assessment was generally perceived negatively, there were some signs of summative assessment being viewed as positive for student learning.

Examples of formative practices included teachers actively providing feedback and extensive coaching of students to involve them in developing self and peer assessment processes. Teachers asked questions that required in-depth replies. They

also worked from a philosophical paradigm where there was choice, which in outdoor education language is referred to as ‘challenge by choice’. This was not about forcing students to engage in some action against their will (Itin, 1996). Instead, they were supported to choose their level of challenge and participation. All these practices helped create a culture that promoted students to become accurate self-assessors, and take responsibility for their own ongoing learning and the continued development of their professional practice. This approach involved teachers and students negotiating the learning challenge through a collaborative model. Students and teachers worked together to agree on the design, delivery, and assessment of the course, providing some ownership of the education process for students and linking teacher practices to assessment *as* learning. In spite of these practices, where many teachers were already involving students in the whole teaching-learning-assessment transaction, teachers appeared unaware of some of the theoretical concepts underlying their assessment practices.

Despite the fact that many teachers could not explicitly articulate the theory, it is pleasing to note that many teachers were already exploring ways to share the education process more with students. They could see the value of students being involved in assessment and the whole process of learning. These few examples demonstrate that teachers have adopted a number of effective teaching methods by employing what Biggs (2003) describes as a good teaching system that aligns teaching methods and assessment to the learning outcomes, and provides learning activities that concentrate on helping students reach course objectives.

Marton and Saljo (1976), also suggest that good teaching comes from teaching methods that entail a deep approach to learning so that students are encouraged to engage with learning and assessment opportunities that generate a genuine desire to complete the task as well as possible. This enables them to maximise the learning opportunity in a meaningful way. Again, the observed teaching practices described above showed teachers achieving this goal. From the practices identified, it appeared that teachers were offering approaches to learning that assisted students to develop their own understanding through a range of learning activities, encouraging them to take responsibility for their own learning. Formative assessment as a process for learning clearly held stronger support than summative assessment, although

summative assessment was reflected by the need for students to meet standards. Summative assessment clearly caused tension for teachers as they tried to balance a supportive learning environment with the need at the same time to provide external credibility for their courses and programmes. Perhaps the biggest hurdle for teachers' involving students in summative assessments was how to maintain standards; a valid concern, given that teachers must maintain course standards.

Teachers were observed asking students on numerous occasions for their self-assessment, and accepting these judgements, probably because they matched their own judgement where the standard was met. Any mis-match of student judgement was approached carefully, with the teacher trying to guide the student to align their judgement with the teacher's. Otherwise the teacher simply overruled the students' self-assessment decision. Observation evidence highlighted that teachers were already working well with their students to improve learning and assessment processes. Teachers who are yet to embrace assessment, which involves students more, perhaps could reconsider their perceived role of the teacher as the sole judge. With some professional development this may not be that difficult, given their existing student centred approach and their focus on learning. Teachers clearly valued self-assessment and regarded this as a stepping-stone towards supporting students becoming independent learners. The benefit of this as Earl (2003a) says, is that self-assessment helps students to be self-motivated and practice using their knowledge to solve problems and make decisions, without waiting for an authority to tell them they have the right answer. It would not be too difficult for teachers to include more assessment by students in the education process.

Assessment Focus

The predominant assessments observed and examples discussed focused assessment on technical skills in outdoor pursuits. Personal competence in pursuits had a high weighting through the observed assessments and in the interviews. At first glance, the focus on technical skills indicated that teachers valued technical skills over other aspects of outdoor education. However, continued observation and some probing questioning, revealed teachers were acutely aware of other aspects that required

assessment. The dilemma for most of teachers was how to assess these other aspects, which were perceived as less tangible.

Despite the difference in assessment between technical skill and other aspects, teachers repeatedly stated that they valued the other aspects like judgement and leadership. Some even expressed their concerns regarding the lack of balance in their assessment practice. Teacher awareness of the need to assess other aspects of outdoor education was high. Yet, such aspects remained largely un-assessed despite teachers saying they were important. This exposes a disjunction between what most teachers say they value and what they assess and this disparity appears to correlate with the perceived difficulty teachers expressed about the challenge of summative assessment.

The lack of assessment of other aspects is, however, not solely caused by the difficulty teachers have with assessment. Two other influences appear to prevent these other aspects being assessed. One factor is the perceived nature of these other aspects, they were considered less tangible, as one teacher said, "*Leadership is still a reasonably nebulous kind of concept" (12: p24),* meaning it is difficult to judge because it is difficult to define. Teaching attributes such as leadership and judgement require a lot of time for teaching and learning and assessment, which was also considered by most to be virtually impossible. Secondly, courses appeared to have no learning outcomes that focused assessment on these other aspects; technical skills dominated the learning outcomes. As one teacher so clearly said: *"The criteria are all based around* (questions like) *can you kayak, can you climb. Yet there is all that other stuff in the curriculum document that says you are constantly assessed on your professional attitude and your attitude to learning. We value these and want to assess them but they don't show up in the assessment criteria, so we cannot report it" (6: p9 & 10).* This example highlights that there was an observed mis-match between individual courses that make up a programme and the graduate profile. For example, graduates are expected to be environmental stewards upon completion of their programme of study, yet there appeared to be minimal assessment of this.

Furthermore, many teachers considered these less tangible aspects far too subjective for them to be assessed, yet teachers were able to identify some of these aspects at

the same time they were assessing skills. Such comments and observations were, however, kept informal and unreported. This lack of other assessment could obviously have some consequences for outdoor education, and raises the question of how end-user groups can be assured that the teachers and instructors delivering outdoor education are suitably qualified and capable. It also raises the possibility that a lack of assessment in some areas of capability may be a contributing factor in recent accidents.

A recent example helps to illustrate this point. In 2008, six students and a teacher drowned while trying to get out of the flooded Mangatepopo Gorge as part of an outdoor education experience. Subsequently the causes of this tragedy were extensively investigated and analysed. Both the coroner's report (Devonport, 2010) and an independent report (Brookes, Smith, & Corkill, 2008), cited the level of experience of the instructor as a possible contributory factor. A key finding from both these reports identified that the experience and judgement of teachers or instructors who work in these environments should be assessed in a range of situations. At the same time, the independent report identified that the wording around policies and competencies is a continuing challenge, as they are open to subjective interpretation. However, as several teachers acknowledged, the writing of objective criteria that clearly define standards is near impossible. This highlights the fact that professional judgement is something that organisations rely on when assessing their staff because they depend on teacher judgements that are made in the field. Nonetheless, as the coroner stated "when an instructor or teacher is provided on a fee-paying basis (even if the provider is a not for profit organisation), it is reasonable to assume that appropriate skills and judgement will be exercised by the provider of the service to avoid serious injury or death" (Devonport, 2010, p. 23). This statement illustrates that, for safety alone, there is a clear need to develop assessment practices within outdoor education to meet a wider range of assessment criteria and possible situations.

If teachers wish to assess these other aspects, an important step would be the inclusion of these facets into explicit learning outcomes. This would enable assessments to occur, making them legitimate, and increasing their visibility, value, and importance. For example, a leadership learning outcome could be: *understand*

principles of leadership with groups and apply these to changing situations in outdoor education environments. This would allow teachers to define what they mean by leadership, to teach theories of leadership, to discuss how, where and when they think it should be applied, to develop behaviour interpretations, to write assessment criteria, and include leadership in their teaching, learning, and formative and summative assessments practices. In order for this to occur, teachers may require assistance in defining these other aspects into written learning outcomes, and in developing assessment criteria. This would also reduce the current perception that these other aspects are overly subjective and un-assessable. The interview data identified that the use of criteria makes all the difference in being able to assess these less tangible aspects.

One teacher did comment on his experience of assessing attitudes. He found it challenging, but it appears his success in doing so has been a result of several teachers in the same institution working together to develop criteria. This showed that other aspects of outdoor education could be effectively assessed. If this is so, perhaps it is teachers' knowledge and perceptions of this form of assessment that may need to be addressed before they can feel comfortable embracing this practice. The perception that assessing other aspects is too intangible and overly subjective appears to be based on a lack of familiarity with assessing these facets, especially when compared to assessment experience in assessing technical skills. In practice, technical skill assessment is probably as subjective, teachers familiarity with technical skill assessment simply leads them to a false sense of security, where they perceive technical assessments to be less risky, less subjective and easier, because that is what they have always done.

Criteria and Professional Judgement in Assessment

The concept of driving assessments towards objectivity and consistency through defined criteria and standards is probably highly desirable, although assessment in reality can never be entirely free of subjectivity. This may be an issue for teachers, who perhaps fail to appreciate the importance of the professional judgement they actually use in making assessment decisions. As Crooks (1988) says, any attempt to rigidly apply specific criteria can interfere with the appropriate use of professional

judgement in the final stage of assessment decisions. While assessing to criteria and standards is common practice, it is apparent that criteria and standards, on their own, can be problematic. One view of assessment under a criterion or standards referenced model requires teachers to base judgement on evidence gathered from a variety of sources against the stated criteria and standards. However, as McDonald, Boud, Francis and Gonczi (n.d.) say, this assumes that criteria and standards can be defined fully, that sufficient evidence can be gathered, that students can achieve and understand these, that teachers can makes sense of differing performances and are able to judge performance consistently.

Examination of criteria revealed that it is difficult to detail criteria and standards so clearly that, no matter how defined the assessment standards are, students and teachers generally seek more detail and clarification. If the criteria are vague, this can lead to a reduction of assessment reliability and validity, and possibly result in inappropriate educational decisions. Reliability concerns can be addressed not by assessment rubrics, but by the use of professional judgement and the use of exemplars. The use of professional judgement, therefore, is a critical ingredient of the assessment process. Teachers become judges of evidence weighing up how well students performances reflect the defined standards (McDonald, et al., n.d.). Nonetheless, criteria probably enable students to better prepare themselves for assessment, even if they are only an indicator of what standard will be required.

Therefore, good practice in assessment probably relies on constructive alignment with criteria stating expected standards along with good judgement being applied. However, judgement is only as good as those that make the decisions; and the use of criteria or standards setting does not guarantee alignment. For example, the criteria used during observations appeared to be no more than general assessment tasks. Words like ‘climb fluidly’ or ‘roll a kayak reliably’ were used to describe a skill movement; however, these words did not clearly explain the standard. Was this one climb, or five? How would a student know what ‘fluidly’ looks like? Was rolling a kayak nine from ten attempts or four from five attempts? To understand the criteria, further explanation would be required. If the criteria and standards are left open, it will probably lead to confusion for teachers and students, and may lead to a variation in assessments standards. This is probably unavoidable, and it would be overly

simplistic to say that teachers need to improve their ability to write criteria and standards more explicitly. This simply denies the contexts and culture that criteria and standards sit in. Wenger (1998) says, that in order to fully understand contexts and culture, teachers need to become a part of a community of practice.

Communities of practice are groups of people who share a concern or a passion for something they do, and learn how to do it better through regular interaction (Wenger, 2007). Assessment is more than the technical knowledge or skill associated with undertaking some task. The group members are involved in a set of relationships over time (Lave & Wenger, 1991) where members value their collective competence and learn from each other, which acts to provide a sense of joint endeavour and identity that underpins professional judgement. Learning and assessment, therefore, becomes more than the mastery of knowledge and skills, but involve a new member (student) entering into a community of practice. Wenger (1998) suggests learning and assessment should not be separated from their communities of practice or the learning becomes irrelevant or abstract. Criteria instead could be better understood in assessment as one aspect of many to help explain what is meant by the assessment standards. Torrance (2007), suggests understanding criteria and the expected standards should occur through the learning process where students' explore and interrogate these and not just accept them. Likewise for teachers', understanding is most likely to happen when they interact with one another.

It has been shown that teachers used criteria to attempt to describe what was required, but in practice they used their professional judgement from a community of practice that supported them to evaluate what they have observed, read, or heard in order to make an assessment decision. This was based on years of experience and expertise, as these teachers suggested. *It is the confidence to know. I have seen thousands of these now and I am an astute judge of when they are working well (6: p31).* Teachers appeared to hold the criteria or standard in their heads and this is commonly known as a 'big picture' view of what competence or meeting the standard looks like. Individual competencies are largely ignored unless too many weaknesses make the big picture unattainable. These competencies are difficult to define, especially in a dynamic environment, like the outdoors. The use of professional judgement, therefore, is an essential component of assessment. It is

possible some teachers apply judgement to small isolated competencies, however, teachers have largely moved away from the checklist-ticking model, to a model where the whole is considered. This practice is endorsed by Biggs (2003), and Sadler (2007) who suggest a valid assessment must be of the total performance, although they suggest there is probably value in assessing components during formative assessment periods.

While the practice of standard setting through criteria has been considered a way to make assessments more explicit and objective, Crooks (1988) also argues any attempt to specify explicit and rigid criteria can prevent teachers from accepting original and innovative work or performance. A rigid approach has the potential to affect development and obstruct appropriate use of professional judgement. Trying to limit subjectivity through prescriptive standards denies the expertise of a teacher. Assessing requires the use of professional judgement, a skill used by teachers everyday to run their classrooms. Professional judgements should not therefore be considered random or arbitrary decisions. As Wenger (1998) says, these verdicts are based on evidence from experienced professionals who practice in a community of professionals that represent collective best practice. To develop more confidence around professional judgement, teachers must be well informed and ensure they benchmark their assessment standards with others in the same field. While this raises the issue of reliability between intra-marking and inter-marker judgements, many teachers were actively involved with NZOIA and discussed standards regularly, moderating and comparing specific examples. Observations indicated their judgements were similar; what amounted to competence was well understood and agreed upon. Judging the benchmark student as a pass or fail grade will always be difficult as exact standards are difficult to define. This helps to clarify why teachers at times struggle with summative assessments.

Nonetheless, it is acknowledged that the use of professional judgement by teachers for assessment decisions usually brings a degree of mistrust to the assessment processes. Crooks (1988) suggests, that the reliability issues that professional judgement poses can be reduced if teachers assess fairly often. This is because the reliability limits of any one assessment are compensated for by the quantity of assessment. However, it is difficult to escape the possibility of error, value driven

decisions and practices which may hinder students' learning and progress. Involving students in assessment as a collaborative processes where there are a variety of tasks would certainly assist these issues and help build trust and acceptance of professional judgement. Teachers grappling with the perception that professional judgement is too subjective to use for decisions undervalue their professional experience. It is important to recognise that professional judgement is a valuable and justifiable tool, which is no more subjective and no less justifiable than criteria that are vague and un-quantified.

The next chapter draws together the outcomes from the findings and the discussion to show how the research questions have been answered.

Chapter 7: Conclusion and Recommendations

The purpose of this Masters study has been to add to the body of knowledge of outdoor education by identifying, documenting and clarifying, current assessment practices of tertiary teachers in outdoor education. It is hoped the results will help provide knowledge and insights that can be used to improve and advance this field of practice.

The research began with a number of perceived issues around assessment in outdoor education, and these helped formulate the research questions of the study. This chapter will draw together the key findings and discussion into a succinct summary, while illustrating how the research questions have been addressed.

Chapter 2 placed the research in context and painted a very complex picture. It highlighted the issue of stakeholders and their varying expectations with regard to outdoor education. One key theme that emerged from several authors (M. Brown & Fraser, 2009; Irwin, 2008; Martin, 2005a) was the unsettled relationship outdoor recreation education has with outdoor education. Irwin (2010) suggested that the use of risk, adventure, and challenge causes outdoor education to be "a discipline arranged around adventurous activities with a lesser emphasis on environmental education" (p. 54). Brown (2008) said outdoor education should move towards place-based education, a form of outdoor education that lowers the level of risk, in order to, Martin (2005a) says, develop and strengthen connections with our environments. However, this research has illustrated that outdoor recreation education remains a highly visible aspect of outdoor education programmes. Teachers and students appear to have a high regard for outdoor recreation education, and therefore it holds a substantial and respected position within their programmes. For teachers, the need for personal competence in the outdoors was unquestioned, and it was clear that outdoor education teachers must have personal capability to offer safe and effective educational experiences with others. However, outdoor education remains a highly contested space. What academics are arguing for and what teachers are practising appear to be somewhat in opposition.

The research revealed that assessment can be used for various purposes; however, the prevailing view has been to measure learning through an emphasis on summative assessment. Policy makers and employers continue to value summative assessment over the promotion of learning. Yet many authors (Biggs, 2003; Dunn, et al., 2004; Race, 2009), argue that the role and purpose of assessment is far broader than maintaining standards and measuring student learning. Assessment should include the diagnosing of student difficulties, measuring student improvement over time, motivating students, judging mastery of essential skills, evaluating teaching, and providing feedback to students. While there are strong arguments that assessment should support learning, broader perspectives of assessment cause tension for stakeholders. Two ideas perhaps summarise this. Firstly, Bryan and Clegg (2006) have suggested that the ability to learn in today's age is perhaps more important than focusing assessment on discrete skills and facts in a hierarchical sequence. Secondly, as Boud (1995) says, "students can, with difficulty, escape from the effects of poor teaching, they cannot (by definition, if they want to graduate) escape the effects of poor assessment" (p. 35). All these differing views about assessment show that it remains highly contested and this is true in New Zealand outdoor education as well.

The findings of the research and the discussion exposed three themes which reflect teacher's practices and illustrate some of their assumptions, values, and processes. To meet the objectives of the study the following discussion encapsulates the key findings presented in the previous chapters and makes evident how the research questions have been addressed.

Perceptions of Assessment

This theme presented assessment as holding contradictory values for teachers. Teacher preferred to be teaching and helping students learn, valuing assessment *for* learning. Assessment *of* learning was found to be a complex and difficult aspect of practice, a process not enjoyed.

Formative assessment was highly valued because it promoted learning and allowed teachers to see the effectiveness of their work. Teachers appeared well practised and highly skilled at assessment *for* learning. Formative assessment emerged as a natural and routine practice, although it was surprising that teachers hardly mentioned this

aspect of practice in the interviews. They appeared to assume that involving students actively in the learning process aided their learning and empowered students to work towards deciding the quality of their own work and developing self-efficacy. While there is support in literature for these ideas (Earl, 2003a; Race, 2009; Sadler, 2007; Torrance, 2007), it seems teachers are not as familiar with the theoretical support backing their practices.

Summative assessment on the other hand, was valued because it maintained standards, but the process did appear to cause teachers considerable anguish as they laboured over summative assessment decisions. The research suggested that this might be a result of inadequate teacher training. Furthermore, teachers routinely used traditional summative assessments that were one-off, high-stakes activities. It seemed this practice had not advanced in accordance with current theory and practice. Sadler (2007), suggested that summative assessment may in some situations benefit from accumulative or progressive assessment judgements, and that formative assessment practices can be used during the learning phase for summative purposes. A change in the way the teachers think about assessment and gather evidence could provide a broader range of evidence, which may well enable teachers to make more reliable assessment judgements and at the same time make summative assessments less complex and laboured. Instead of teachers viewing summative assessment in a negative context, assessment could become a part of their day-to-day practice to motivate students, reduce the complexity and pressure of one-off, high-stake assessments, and legitimise the formative judgements, they already use, for summative purposes.

Assessment Focus

The data revealed that there was an over-emphasis on assessing technical skills and a lack of assessment of other essential aspects, for example, leadership, judgement, and decision-making.

Teachers themselves were highly capable in technical skills, which helped to explain, perhaps, the high value they placed on personal capability in outdoor recreation. Students were expected to be able to perform to a reasonable standard in this area. Teachers seemed to assume that assessing technical skills was a key aspect for

providing safety in outdoor education. It is difficult to argue against this logic, given recent accidents within New Zealand. However, other key skills, knowledge, and attitudes remain outside the range of explicit assessment. This highlights an issue for outdoor education teachers. Assessing technical skills capability in isolation from other essential aspects such as leadership and decision-making competence cannot be considered an overall judgement of a student's capability in real-time situations. The issue for teachers is how to assess these other aspects, as there appeared a genuine willingness to do so, in spite of the fact that other aspects were perceived as less tangible. For example, the assessment of leadership was considered more subjective, and teachers therefore tended to avoid such an area. This means that there is a probable imbalance in the curriculum design, content, and assessment focus.

Alternative pedagogical practices, as Humberstone (2000) and Preston (2001) point out, can only be reached through paradigm shifts which require a willingness on the part of teachers to critically reflect on their practices. Current practice of outdoor education, according to Humberstone (2000), remains locked into traditional male paradigms, which Green (1994) has described as an emphasis on physical strength, speed and technical expertise. Lugg (2003) and Preston (2001), suggest teachers have a moral responsibility to challenge dominant discourses such as this, to help develop new knowledge and practices which may improve the broader field or profession.

Assessment Criteria and Professional Judgement in Assessment

Teachers appeared to view the application of assessment criteria as being incongruous with the use of professional judgement. The underlying assumption was that criteria define the standard and are supposed to be objective, while professional judgement uses experience to make a decision about standards and is overly subjective. Because of this perception, the use of preset criteria was perceived to have greater credibility and validity, as assessment decisions were based on explicit standards. Teachers tended to deny the place of professional judgement as an assessment tool, as it was considered less reliable, more contestable, and overly subjective. Teachers' trust in preset assessment criteria is perhaps somewhat misplaced as in practice the assessment criteria used were little more than assessment tasks; standards were not well defined. This meant that teachers were compelled to use professional judgement in order to make assessment decisions.

This exposes an assumption, that the use of criteria makes assessments more explicit, objective, and consistent and free of subjectivity. This is challenged by Crooks (1988), who points out that specifying explicit and rigid criteria can prevent teachers from accepting original and innovative work or performances and can affect the appropriate use of professional judgement. He argues that attempts at limiting subjectivity through prescriptive standards denies the legitimate role of a teacher. In fact, professional judgement by teachers should not be considered random or arbitrary decisions, but an essential aspect of an assessment process where judgements are based on professional experience and expertise. Woolf (2004) supports this notion by saying "professional and academic judgement is central to determining grades" (p. 487). Professional judgement is a part of what Wenger (1998) would call a community of practice, where group members who are experts make decisions based on evidence from experience. Teachers are experts who are also members of a collective of practice and their judgements can, therefore, be trusted and relied upon.

The use of criteria and professional judgement should not be considered mutually exclusive. Instead, they act to support each other to help deliver fair and objective (as possible) assessment decisions. Any reliability concerns can be addressed not by strict assessment rubrics, but by the use of professional judgement and the use of exemplars. Professional judgement is no more subjective and unjustifiable as an assessment tool than the application of criteria, especially when the criteria used are vague and un-quantified. Teachers grappling with the concept of professional judgement being too subjective, undervalue their professional experience (Crooks, 1988).

Revisiting the research questions

The following summarises the key outcomes of the research, which have been discussed through the discussion and conclusion chapters.

What are the assessment practices of tertiary teachers delivering outdoor education qualifications throughout New Zealand?

1. Emphasis was placed on practical assessment.

2. Assessments aspects that were assessed were through a holistic approach.
3. Formative assessment practices were prevalent and highly valued for the purpose of improving learning.
4. Learning through direct experiences where consequences could be powerful and provide instant feedback was common practice.
5. Assessment processes that supported self assessment were common and valued as a path to continued professional practice.

What are the assumptions underpinning teacher assessment practices?

1. The assessment of less tangible aspects of outdoor education practice is too subjective to enable both valid and reliable assessment.
2. Students who perform well in technical skills can develop the other skills and knowledge as their experience expands.
3. Outdoor recreation is a fundamental aspect of outdoor education.

What assessment processes do they use and why do they use them?

1. Teachers routinely and naturally practice formative assessment to assist students in their learning.
2. Summative assessment is seen as separate to formative assessment and as such is a process that is usually applied only at the end of a course of study.
3. Formative assessment judgements stand alone and do not appear to contribute to summative results. This is possibly due to perspectives on assessment where summative assessment is through one-off, high-stake events.

What values are teachers communicating through their assessment practices?

1. Technical skills in outdoor pursuits are a key aspect of outdoor education programmes.
2. Students must be able to perform at high levels of personal competence to be an effective outdoor educator.

Recommendations

There are many recommendations that could come from this research and I recognise that readers may well have identified different recommendations to those below. For improved practice, the key recommendations I would make are as follows:

1. That tertiary outdoor education teachers explore ways to include the assessment of other aspects of being an effective outdoor educator than technical skills. This probably means, for example, assessing students as they run clients through outdoor education experiences. In such situations qualities like judgement and decision-making are required, and can be assessed through real situations.

2. That tertiary outdoor education teachers investigate and trial assessment practices which include the accumulation of evidence from day-to-day teaching in formative situations to be used for summative assessment decisions.

3. That tertiary outdoor education teachers reflect on their formal assessment education and training experience and if they identify gaps in their education commit to professional development.

4. That tertiary outdoor educators develop a professional development programme as a community of practice to support the making and justifying of professional judgements in assessment.

5. Assessment criteria, in some cases, should be phrased in more specific terms, and it should be explained to students and teachers how these are to be used and what functions they serve.

Recommendations for future research

There are many directions that future research might take, and as a qualitative researcher, I recognise that readers will have most certainly identified opportunities that I have not considered. However, the discussions that have taken place have signalled several key opportunities for further research relating to assessment practices in outdoor education. These key opportunities for further research are:

1. Exploration of the effects of tertiary teachers receiving assessment education and the outcomes of this on their assessment practices, the reliability of their judgements and their confidence in those judgements.

2. Examination of ways to shift outdoor education training away from focusing predominately on assessing technical skills and knowledge to a broader consideration of human factors that make outdoor educators safe and effective teachers.

3. Research effective professional judgement practices in outdoor education.

Concluding thoughts

This Masters degree has spanned some six years and this research some two and a half years: it is interesting to reflect upon my own learning.

Often when people are at home, they dream about and crave adventures away, and when they are on those adventures, they dream about and crave to be home. This research journey has surely been an adventure, one that has challenged me, one that has held my attention most of the time, a journey that I have mostly enjoyed.

It has created many more questions. I have pondered the validity of 'interpretive' research. In trying to resolve this, I thought about the contrasting perspectives of scientific research, which seems to pride itself on working with 'facts and figures', as if that data is more objective. However, with more thought, I have realised that even facts and figures have to be interpreted. Quantitative research also includes aspects of interpretation. This made me wonder, why 'quantitative' research appeared to have

greater prestige. Schon (1987) has offered some insight. He points out that just as 'school knowledge' is more highly esteemed than 'knowing in practice/action'; qualitative research also seems to suffer from the same prejudice. The result is that some consider it less valuable. It is Schon's thinking that helped me to see greater value in interpretive approaches, and highlighted how narrow an individual view of the world can be.

Reflecting on the methodology used

As detailed above, this study fell within the qualitative research paradigm. Using the illuminative evaluation methodology to gain insights into the assessment practices of tertiary outdoor education teachers in New Zealand has entailed studying teachers assessing at a variety of different teaching locations around New Zealand. By adopting the illuminative approach, this study was able to provide a rich and detailed description of teacher assessment practices. The research approach meant a variety of data-gathering methods could be used. This resulted in an abundance of data and showed that this approach was a highly fruitful way to gather data about a complex process. While the descriptive and interpretive nature of this approach inevitably involved a degree of subjectivity, one standout feature of this methodology was the way it provided multiple opportunities to progressively re-focus and cross-check the data. Each stage of the data gathering phase enabled a great deal of cross checking, which ensured a high degree of validity. This approach also provided a strategy for dealing with limited amounts of research in this area.

If the research had been approached using predetermined, standard evaluation criteria and instruments, it is doubtful whether many of these insights would have been gained. However, this approach also presented a number of difficulties. The data that had to be gathered came from a number of different learning locations throughout New Zealand and this involved costs. Thankfully, these expenses were mostly met through a Massey University research grant. Also, because of the extensive data gathered, it took a considerable time to analyse and process the data. In short, the greatest disadvantages of this methodology were that it was both time-consuming and somewhat expensive to conduct. However, given the depth of insight provided by the methodology, its benefits I believe outweigh these negative factors. What

eventuated from the illuminative evaluation approach was a very useful framework for the research.

Final comments…

Research is a funny thing, just when you think it is finished more thoughts, perspectives, issues, and ideas present themselves, so while this is the end for now, it is not the end of what could be researched in the future. The topic lives on…

There is no doubt in my mind that this research study has been a wonderful opportunity and experience. It has provided me with a chance to become really engaged with the topic and to reflect not only on my practices, but to put ideas on paper.

References

Adams, L. (2007). New Zealand Outdoor Leader Standards and Qualifications Alignment Project: Phase II. Retrieved from http://www.outdoorsnz.org.nz/cms_display.php?sn=43&pg=1307&st=1

Allin, L. (2000). Women into outdoor education: Negotiating a male-gendered space - issues of physicality. In B. Humberstone (Ed.), *Her outdoors, challenge and adventure in gendered open spaces* (Vol. 66). Brighton, UK: Leisure Studies Association.

Anderson, L., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives*. New York, USA: Longman.

Angelo, T. (2003). Doing assessment as if learning matters most. *American Association for Higher Education*, 1-10.

Assessment Reform Group (2002). *Assessment is for learning: 10 principles.* Retrieved Oct, 2010, from www.assessment-reform-group.org/CIE3.PDF

Astin, A. (1993). *Assessment for excellence*. Phoenix, USA: The Oryx Press.

Atherton, J. (2009). Learning and teaching: Deep and surface learning. Retrieved April 2010, from http://www.learningandteaching.info/learning/deepsurf.htm

Atkinson, M. (2009). Are they prepared? Training opportunities and qualifications for teachers of outdoor and adventure education. Unpublished Dissertation. Lincoln University.

Austin, T. (2005). Knowledge capture, knowledge rendering and knowledge use in a metropolitan fire service. In A. Hope & P. Oliver (Eds.), *Risk, education and culture*. Hampshire, UK: Ashgate Publishing Limited.

Barnett, R. (1994). *The limits of competence*. Buckingham, UK: Society for Research into Higher Education & Open University Press.

Baume, D., Yorke, M., & Coggey, M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment & Evaluation in Higher Education, 29*(4), 451-477.

Beedie, P. (1995/6). Where are the risk takers in outdoor education? *Journal of Adventure Education and Outdoor Leadership, 12*(4), 10-13.

Biggs, J. (2003). *Teaching for quality learning at university*. Berkshire, UK: Open University Press.

Black, P. (1998). *Testing: friend or foe?* London, UK: Falmer Press.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, UK: Open University Press.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education, 21*(2), 119-132.

Bloom, B. (1956). *Taxonomy of educational objectives, the classification of educational goals*. New York, USA: McKay.

Boud, D. (1995). Assessment and learning: Contradictory or complementary? Assessment for learning in higher education. In P. Knight (Ed.), (pp. 35-38). London, UK: Kogan Page.

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

Boud, D., & Falchikov, N. (Eds.). (2007). *Rethinking assessment in higher education*. New York, USA: Routledge.

Boyes, M. (2004). The maintenance of quality in preparation of outdoor education teachers. *New Zealand Journal of Outdoor Education, 1*(4), 82-98.

Boyes, M., & Zink, R. (2006). The nature and scope of outdoor education in New Zealand schools. *Australian Journal of Outdoor Education, 10*(1), 11-21.

Broadfoot, P. (1996). *Education, assessment and society*. Buckingham, UK: Open University Press.

Brookes, A. (1991). *Thoughts on the conference theme.* Paper presented at the Proceedings of the 7th National Outdoor Education Conference, Monash University, Frankston, Australia.

Brookes, A. (2003). A critique of Neo-Hahnian outdoor education theory. Part one: Challenges to the concept of "character building". *Journal of Adventure Education and Outdoor Learning, 3*(1), 49-62.

Brookes, A., Smith, M., & Corkill, B. (2008). *Independent report to the trustees of the Sir Edmund Hillary Outdoor Pursuit Centre of New Zealand. Mangatepopo Gorge Incident*.

Brown, M. (2008). Outdoor education: Opportunities provided by a place based approach. *New Zealand Journal of Outdoor Education, 2*(3), 7-25.

Brown, M., & Fraser, D. (2009). Re-evaluating risk and exploring educational alternatives. *Journal of Adventure Education and Outdoor Learning*, *9*(1), 61-77.

Brown, S., & Glasner, A. (Eds.). (1999). *Assessment matters in higher education*. Buckingham, UK: SRHE and Open University Press.

Bruniges, M. (2007). Teacher professional judgement in teaching and learning decisions. In R. Maclean (Ed.), *Learning and teaching for the twenty first century* (pp. 233-254). New York, USA: Springer.

Bryan, C., & Clegg, K. (Eds.). (2006). *Innovative assessment in higher education*. Oxon, UK: Routledge.

Buell, L. (1983). Outdoor leadership competency: A manual for self-assessment and staff evaluation. *Environmental Awareness Publications*.

Burden, B. (2008). Illuminative evaluation. In B. Kelly, L. Woolfson & J. Boyle (Eds.), *Frameworks for practice in educational psychology*. London, UK: Jessica Kingsley Publishers.

Burns, N., & Grove, S. (1993). *The practice of nursing research: Conduct, critique, and utilization*. Missouri, USA: Elsevier/Saunders.

Carpenter, D., Gunn, D., & Cant, M. (2009). *Memorandum of understanding between Skills Active, NZOIA & MSC.* Retrieved December, 2009, from http://www.nzoia.org.nz/assets/images/Outdoor%20MOU%20Skills%20Active%20MSC%20NZOIA%20Dec%202009.pdf

Carter, M. (2000). Developing confidence in women working outdoors: An exploration of self confidence and competence in women employed in adventure recreation. In B. Humberstone (Ed.), *Her outdoors, challenge and adventure in gendered open spaces* (Vol. 66). Brighton, UK: Leisure Studies Association.

Chambers, M. (1988). Curriculum evaluation: An approach towards appraising a post-basic psychiatric nursing course. *Journal of Advanced Nursing, 13*(3), 330-340.

Chapman, H. (1999). Some important limitations of competency-based education with respect to nurse education: An Australian perspective. *Nurse Education Today 19*(2), 129-125.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London, UK: Routledge Falmer.

Crooks, T. (1988). Assessing student performance. *Higher Education Research and Development Society of Australasia, 8*.

Cross, V., Hicks, C., & Barwell, F. (2001). Exploring the Gap Between Evidence and Judgement: Using video vignettes for practice-based assessment of physiotherapy undergraduates. *Assessment & Evaluation in Higher Education, 26*(3), 189-212.

DeLuca, C., & Klinger, D. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419-438.

Denzin, N., & Lincoln, Y. (2008). The discipline and practice of qualitative research. In N. Denzin & Y. Lincoln (Eds.), *The landscape of qualitative research*. California, USA: Sage.

Department of Labour (2010). *Review of risk management and safety in the adventure and outdoor commercial sectors in New Zealand.* Retrieved January, 2011, from http://www.dol.govt.nz/consultation/adventure-tourism/final-report/01.asp

Devonport, C. J. (2010). *Coroner's report regarding Mangatepopo Stream tragedy on 15 April 2008*. Auckland, New Zealand: Coroners Court.

Dewey, J. (1938/1997). *Experience & Education*. New York, USA: Simon & Schuster.

Disraeli, B. (2006). Benefits and values of outdoor recreation. In R. Jensen & S. Guthrie (Eds.), *Outdoor recreation in America* (6th ed.). Champaign, USA: Human Kinetics.

Dochy, F. (2009). The edumetric quality of new modes of assessment: Some issues and prospects. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education.* Wollongong, Australia: Springer.

Draper, S. (1997). Observing, measuring, or evaluating courseware: A conceptual introduction. Retrieved 2009, from http://www.icbl.hw.ac.uk/ltdi/implementing-it/measure.htm

Dumble, R. (2002). *When is a qualification not sufficient? When a law court judge says so!* Paper presented at the New Zealand conference on outdoors risk management, Sir Edmund Hillary Outdoor Pursuit Centre, Turangi, NZ.

Dunn, L., Morgan, C., O'Reilly, M., & Parry, S. (2004). *The student assessment handbook*. London, UK: Routledge Falmer.

Earl, L. (2003a). *Assessment as learning*. California, USA: Corwin Press.

Earl, L. (2003b). *Using classroom assessment to maximise student learning*. California, USA: Corwin Press.

Earl, L., & Katz, S. (2006). *Rethinking classroom assessment with purpose in mind*. Winnipeg, Canada: Manitoba Education, Citizenship and Youth.

Ecclestone, K. (2007a). Commitment, compliance and comfort zones: The effects of formative assessment on vocational educational students' careers. *Assessment in Education: Principles, Policy and Practice, 14*(3), 315-333.

Ecclestone, K. (2007b). Students' experiences in post school qualifications. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education*. London, UK: Routledge.

Fetterman, D. (1989). *Ethnography step by step*. Newbury Park, California, USA: Sage.

Flick, U. (2002). *An introduction to qualitative research*. London, UK: Sage.

Gephart, R. (1999). Paradigms and research methods. *Research Methods Forum, 4*.

Gipps, C. (1994). *Beyond testing*. London, UK: Falmer.

Green, M. (1994). A study of some of the influences affecting the learning of tertiary women involved in outdoor education. Unpublished Master's thesis. Deakin University.

Haddock, C. (1999). Teachers talking about their epics: Near misses in outdoor education. Unpublished Master's thesis. Lincoln University.

Haddock, C. (2004). *Outdoor safety: Risk management for outdoor leaders*. Wellington, NZ: New Zealand Mountain Safety Council Inc.

Hamilton, D. (1976). *Curriculum evaluation*. London, UK: Open Books.

Harvey, J., & Gunn, C. (2000). Online Tutoring e-Book: Chapter 5: Evaluation. In C. Higgison (Eds.) Available from http://otis.scotcit.ac.uk/onlinebook/otisT505.htm

Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. In L. Meyer, S. Davidson, H. Anderson, R. Fletcher, P. Johnston & M. Rees (Eds.), *Tertiary assessment & higher education student outcomes: Policy & research*. Wellington, NZ: Ako Aotearoa.

Health and Safety in Employment Act (1992). Retrieved January, 2010, from http://www.legislation.govt.nz/act/results.aspx?search=ts_act_health+and+safety_resel&p=1

Hill, J. (1998). Rethinking assessment: The challenging issues for schools and teachers. Unpublished Master's Thesis. Massey University.

Holt, L. (1974). Identification of the competencies needed by the classroom teacher in programms of resident outdoor education. Ohio University.

Hovelynck, J. (2001). Beyond didactics: A reconnaissance of experiential learning. *Australian Journal of Outdoor Education, 6*(1), 4-14.

Humberstone, B. (2000). The 'outdoor industry' as social and educational phenoma: Gender and outdoor adventure/education. *Journal of Adventure Education and Outdoor Learning, 1*(1), 21-35.

Hume, A., & Coll, R. (2009). Assessment of learning, for learning, and as learning: New Zealand case studies. *Assessment in Education: Principles, Policy and Practice, 16*(3), 269-290

Irwin, D. (2008). Weaving the threads: Challenges encountered while educating for sustainability in outdoor education. *New Zealand Journal of Outdoor Education, 2*(3), 36-53.

Irwin, D. (2010). *Weaving the treads of education for sustainability and outdoor education.* Unpublished Doctor of Philosophy, Canterbury University, Christchurch, NZ.

Itin, C. (1996). Challenge by choice as a professional enabling. *Rocky Mountain Region of the Association for Experiential Education, 6*.

Jones, A. (1999). The place of judgement in competency-based assessment. *Journal of Vocational Education and Training, 51*(1), 145-160.

Jones, A. (2006). VET professionals making good judgements. In S. McKenna & J. Mitchell (Eds.), *Professional judgment in vocational education and training: A set of resources*. Canberra, Australia: n.p.

Jones, M. (2005). Is there a disconnection between outdoor education practice and outdoor education principles? *Out and About, 13*, 29-32.

Joughin, G. (2009a). Assessment, learning and judgement in higher education: A critical review. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education*. Wollongong, Australia: Springer.

Joughin, G. (2009b). Introducing: Refocusing assessment. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education*. Wollongong, Australia: Springer.

Kiewa, J. (2001). Stepping around things: Gender relationships in climbing. *Australian Journal of Outdoor Education, 5*(2), 4-12.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate periheral participation*. Cambridge, UK: University of Cambridge Press.

Law, B., & McConnell, R. (2007). Aims of environmental education. Retrieved from http://www.tki.org.nz/r/environ_ed/guidelines/aims_enviro_e.php

Light, G., Cox, R., & Calkins, S. (2009). *Learning and teaching in higher education: The reflective professional. (2nd Ed)*. London, UK: Sage.

Luckner, J., & Nadler, R. (1997). *Processing the experience (2nd Ed)*. California, USA: Kendall/Hunt Publishing Company.

Lugg, A. (2003). Women's experience of outdoor education: Still trying to be 'one of the boys'? In B. Humberstone, B. Heather & R. Kaye (Eds.), *Whose Journeys? The Outdoors and Adventure as Social and Cultural Phenomena*. Sydney, Australia: The Institute for Outdoor Learning.

Lynch, P. (1993). Outdoor education - semantics and definitions. *Outdoor Education Research Network Newsletter*.

Lynch, P. (2006). *Camping in the curriculum*. Auckland, NZ: PKL Publications.

Marshall, B., & Drummond, M. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education, 21*(2), 133-149.

Martin, P. (2005a). Human to nature relationships through outdoor education. In T. Gray & B. Hayllar (Eds.), *Outdoor and experiential learning: Views from the top*. Dunedin, NZ: Otago University Press.

Martin, P. (2005b). *Outdoor leadership competence in academic pathways: Or how can 4-year trained teachers of outdoor education demonstrate competence?* Paper presented at the 14th National Outdoor Education Conference, Australia.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning II: Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology, 46*(2), 115-127.

Massey University (2009). *Code of ethical conduct for research, teaching, and evaluations involving human participants.* Retrieved June, 2009, from http://www.massey.ac.nz/massey/fms/Human%20Ethics/Documents/MUHEC%20Code%202010.pdf

McDonald, R., Boud, D., Francis, J., & Gonczi, A. (n.d.). New perspectives on assessment. *Studies in Technical and Vocational Education*(4).

McMillan, H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research and Evaluation, 7*(8), 1-5.

Miles, J., & Priest, S. (1990). *Adventure education*. Pennsylvania, USA: Venture.

Ministry of Education (1999). *Health and physical education.* Retrieved January, 2010, from http://www.tki.org.nz/r/health/curriculum/statement/hpe_statement.pdf

Ministry of Education (2009). *EOTC guidelines. Bringing the curriculum alive.* Retrieved January, 2010, from http://eotc.tki.org.nz/EOTC-home/EOTC-Guidelines

Mortlock, C. (1984). *The adventure alternative*. Milnthorpe, UK: Cicerone Press.

Moyle, S. (2004). Improving integrity by removing equivocation: A critque of NZOIA Rock 1 syllabus. *New Zealand Journal of Outdoor Education, 1*(4), 73-81.

Munge, B. (2007). *A study of Australian employers' perceptions of graduates from the Bachelor of Arts (Outdoor Education) at La Trobe University.* Paper presented at the 15th National Outdoor Education Conference, Victoria, Australia.

Murphy, R. (2006). Evaluating new priorities for assessment in higher education. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education*. London, UK: Routledge.

Mussweiler, T. (2003). Comparison processes in social judgement: Mechanisms and consequences. *Psychological Review, 110*(3), 472-489.

Neill, J. (1997). *Gender: How does it affect the outdoor education experience?* Paper presented at the Catalysts for Change: 10th National Outdoor Education Conference, Sydney, Australia.

Neill, J. (n.d.). What is outdoor education? Retrieved from http://www.wilderdom.com/definitions/definitions.html#DefOAE

Parlett, M., & Hamilton, D. (1972). Evaluation as illumination: A new approach to the study of innovation programs. Centre for Research in the Educational Sciences, University of Edinburgh.

Parlett, M., & Hamilton, D. (1976). Evaluation as illumination. In D. Tawney (Ed.), *Curriculum evaluation today: Trends and implications*. London, UK: Schools Council Research Studies.

Patton, M. (2002). *Qualitative research and evaluation methods.* (3rd ed.). California, USA: Sage.

Pidgeon, N., & Henwood, K. (2004). Grounded theory. In M. Hardy & A. Bryman (Eds.), *Handbook of data analysis*. London, UK: Sage.

Preston, L. (2001). Seeing the unseen: A rock climbing experience. *Australian Journal of Outdoor Education, 5*(2).

Priest, S., & Gass, M. (1997). *Effective leadership in adventure programming*. Illinois, USA: Human Kinetics.

Race, P. (2009). Towards assessment as learning. *The All Ireland Journal of Teaching and Learning in Higher Education, 1*(1).

Ramsden, P. (1992). *Learning to teach in higher education*. London, UK: Routledge.

Rensburg, E. (2007). *An illuminative evaluation of the workplace learning component of UNISA's diploma in animal health.* Witwatersrand, Johannesburg.

Sadler, R. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education: Principles, Policy and Practice, 14*(3), 387-392.

Schon, D. (1987). *Educating the reflective practitioner: Towards a new design for teaching and learning in the professions*. San Francisco, USA: Jossey-Bass.

Schwartz, P. (2002). Gain without pain. In P. Schwartz & G. Webb (Eds.), *Assessment: Case studies, experience and practice from higher education*. London, UK: Biddles Ltd.

Seaman, J., & Coppens, A. (2006). Repertoire of practice: Reconceptualising instructor competency in contemporary adventure education. *Journal of Adventure Education and Outdoor Learning, 6*(1), 25-37.

Smith, R. (1999). Paths of judgement: The revival of practical wisdom. *Educational Philosophy and Theory, 31*(3), 327-340.

SPARC (2007). *Creating a healthy state for outdoor recreation in New Zealand - Discussion document*. Wellington, NZ: Sport and Recreation New Zealand.

SPARC (2009a). *Outdoor recreation strategy 2009-2015.* Retrieved January, 2010, from

http://www.sparc.org.nz/Documents/Publications/Outdoor%20Recreation%20Strategy%202009-15/Outdoor_Recreation_Strategy_2009-15.pdf

SPARC (2009b). *Sir Edmund Hillary Outdoor Recreation Council.* Retrieved January, 2010, from http://www.sparc.org.nz/about-sparc/people-at-sparc/sir-edmund-hillary-outdoor-recreation-council

Speck, B. (1998). Unveiling some of the mystery of professional judgement in classroom assessment. *New Directions for Teaching and Learning, 74*, 17-31.

Stetson, C. (n.d.). An essay on Kurt Hahn, founder of Outward Bound. Retrieved from http://www.kurthahn.org/writings/stet.pdf

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta, 83*(10), 758-765.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques (2nd ed.)*. London, UK: Sage.

Swiderski, M. (1984). *Outdoor leadership competencies: A research study surveying outdoor leaders.* Paper presented at the American Camping Association Convention, San Diego, USA.

Tomlinson, A., Ravenscroft, N., Wheaton, B., & Gilchrist, P. (2005). *Lifestyle sports and national sport policy: An agenda for research*: Report to Sport England.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy and Practice, 14*(3), 281-294.

Tripp, D. (1993). *Critical incidents in teaching: Developing professional judgement.* New York, USA: Routledge.

Volante, L. (2009). *Assessment of, for, and as learning within schools: Implications for transforming classroom practice.* Paper presented at the International Congress for School Effectiveness and Improvement, Vancouver, British Columbia, Canada.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 749-770.

Ward, B. (2007). *Position paper: Outdoor recreation review*. Wellington, NZ: New Zealand Recreation Association.

Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science, 33*, 283-309.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge, UK: Cambridge University Press.

Wenger, E. (2007). Communities of practice: A brief introduction. Retrieved January, 2011, from http://www.ewenger.com/theory

Wheaton, B. (Ed.). (2004). *Understanding lifestyle sports*. New York, USA: Routledge.

Woolf, H. (2004). Assessment criteria: Reflections on current practices. *Assessment and Evaluation in Higher Education, 29*(4), 479-493.

Yandell, J., & Turvey, A. (2007). Standards or communities of practice? Competing models of workplace learning and development. *British Educational Research Journal, 33*(4).

Zink, R., & Leberman, S. (2001). Risking a debate - redefining risk and risk management: A New Zealand case study. *The Journal of Experiential Education, 24*(1), 50-57.

Appendices

Appendix 1: Research Questionnaire

Question 1

Why do you assess?

In my teaching/instructing I assess for (what reasons)…

Question 2

What are the five most important aspects within outdoor education that you assess?

Question 3

What makes these aspects important?

Question 4

How do you assess these important aspects?

Question 5

When during a course do you perform assessments?

	Always	Frequently / Often	Sometimes	Seldom	Never
formatively at the beginning of each course					
formatively throughout the course					
summatively at the end of the course					
summatively throughout the course					
summatively when the student is ready					
summatively when it is timetabled					
never					
other					

Question 6

In your teaching and assessment practices, do you include any of the following?

	Always	Frequently / Often	Sometimes	Seldom	Never
marking against implicit criteria					
marking against explicit criteria					
matching assessments to learning outcomes					
pre-moderation					
post-moderation					
double marking					
blind marking					
second marker if a fail grade is given					
the use of professional judgement					
checking of assessment tasks to ensure there are no inherent biases					
a clear alignment between what is taught and learnt, and the knowledge and skills assessed					
the use of authentic and real tasks					

Question 7

How is the learning assessed? What methods / modes do you use?

In my teaching I assess learning through…

Question 8

In your teaching who is involved in making assessment judgements?

	Always	Frequently / Often	Sometimes	Seldom	Never
self					
peers					
teacher					
employer					
external assessor					
combination of above					
other					

Question 9

What aspects of assessments work well for you?

Question 10

What problems do you experience in assessment?

Question 11

Any other comments you would like to make on assessment?

Question 12

What industry qualifications do you hold if any?

Question 13

Demographic information

What is your...

a. gender
b. age
c. educational qualifications, if any
d. industry qualifications, if any
e. teaching qualification, if any
f. number of years teaching/instructing in higher education
g. overall number of years instructing in outdoor education

Appendix 2: Semi-structured Interview Questions

Introduction sheet – purpose; interview recorded; confidential; right to withdraw; recorder turned off; consent form signed.

Biographical

Can you tell me about: your personal experience in Outdoor Ed and
your teaching/Instructing experience?
subject areas do you teach?
your years of teaching experience?
formal tertiary qualifications you have?
formal industry qualifications you have?

Why assess?

What springs to mind when you think about assessment?
When you assess what are you trying to achieve? Why?
What is the relationship between assessment and learning in your teaching?
When assessing students, what factors do you consider?
What motivates your assessment methodology?

What you assess?

Describe some of your typical assessments?
What is it you are generally assessing for? Why these aspects?
In your teaching and assessment programme, what is the balance between the academic understanding/thinking aspects and the applied practice/performance skills aspects?
How do the likes of judgement, decision making, problem solving, attitudes etc fit into your assessments? Why? How are these assessed / represented?

How do you assess?

<u>Standards:</u>

How do you develop assessment material, benchmarks, exemplars, criteria etc?
What strategies do you use to ensure that students understand the criteria that will be applied in assessing their work?

How do you know a student meets a benchmark or standard?
How do you make these judgements? What are these judgements based on?
What do you do around the involvement of students in assessment decisions?
How do you make sense of inconsistent performance / or differing standards / evidence?
What do you do if you have different student ability levels? genders? ethnicities?

Tools / Approaches:

What variety of assessment methods do you use? Why do you use these? For what reasons/gains/purposes?
How would you describe your assessment practices in general: holistic or reductionist? Why do you assess like this?
What do you do when a students' work is below par or above?
What do you to manage perceptions of objectivity, subjectivity, validity, reliability and fairness?
How are we assured of inter-rater judgements?

When you assess?

Where do assessments fit into your teaching course? – summative at the end of a course?, formative throughout?, integrated? Why?
How are your assessments timed? Why are they timed like this?
How do you involve students in decision making about when assessment should be carried out?
What opportunities do students get to show improvement or reassessment?

Challenges

What are some of your key issues in assessment?

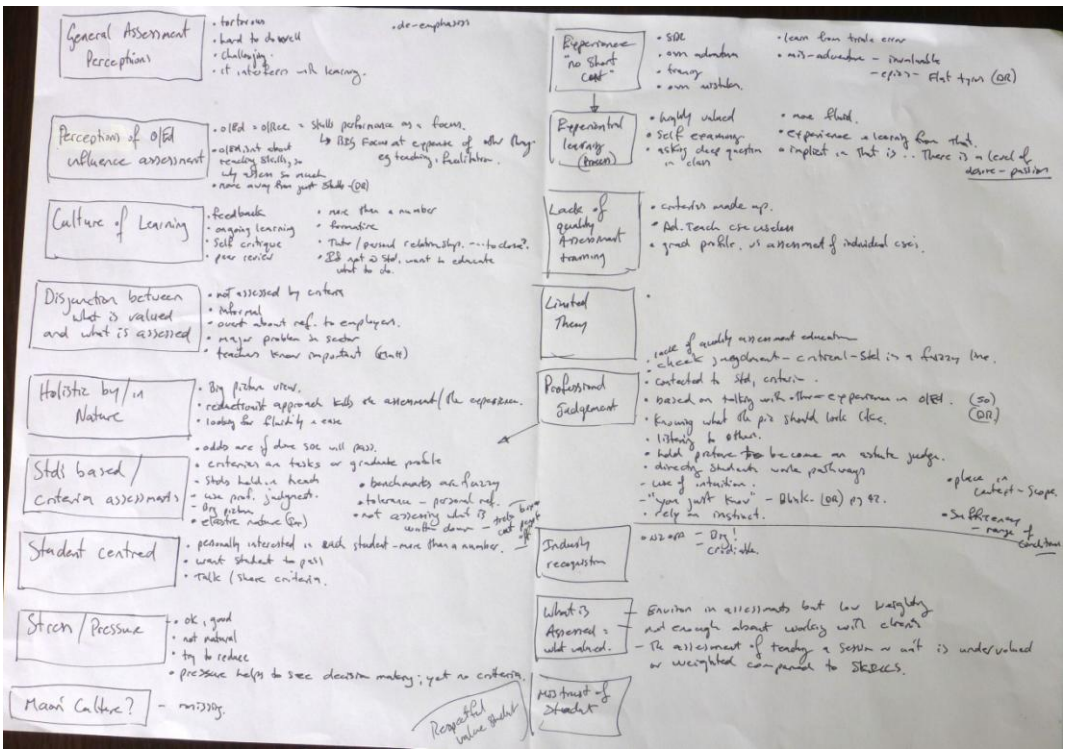
Professional development / Understanding

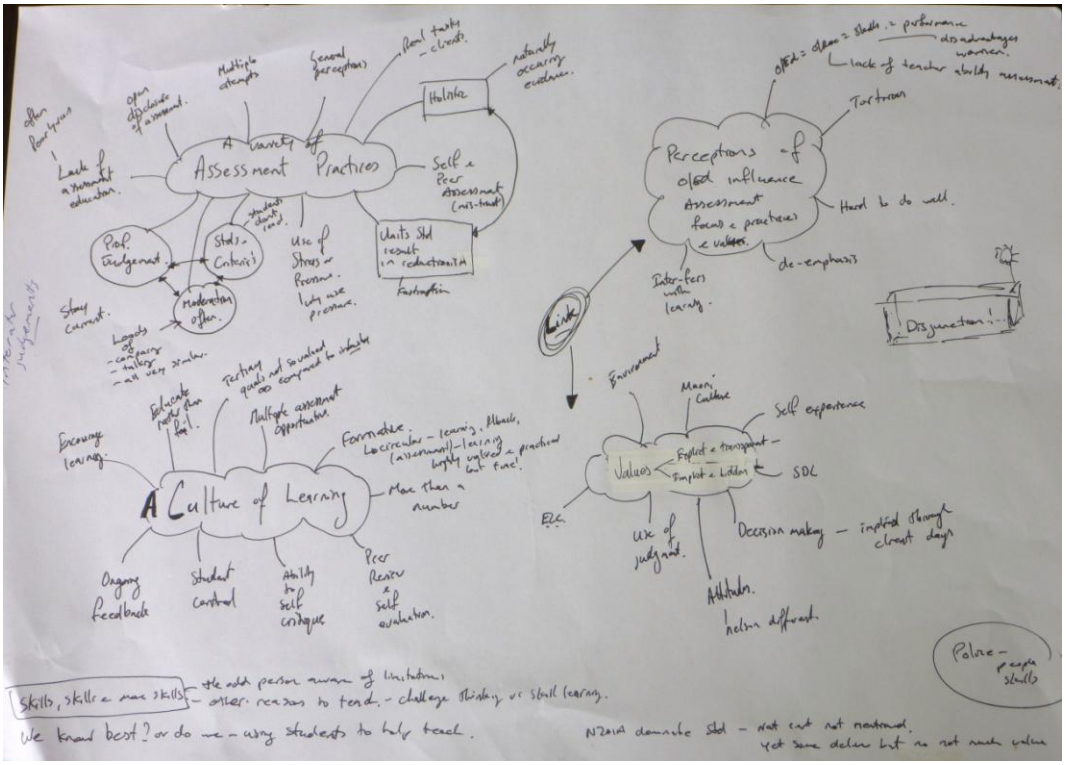
How have you developed your knowledge and assessment strategies/practices about assessment?
What do you do around moderation?
What do you consider to be the most challenging aspect of being a teacher involved in assessment?

What do you consider has been the most rewarding aspect of being a teacher involved in assessment?

Are there any other comments you would like to make about your assessment practices?

Appendix 3: Theme building





Appendix 4: Research Information Sheet

Research Question
What are the assessment practices of teachers of Outdoor Education qualifications in higher education programmes throughout New Zealand?
Sub-questions
1. What are the assumptions underpinning teacher assessment practices?
2. What assessment processes do they use and why do they use these?
3. What values are assessors communicating through their assessment practices?

Name of Researcher: Steve Chapman

Contact:

Information
This research is part of my study towards a Master degree in Education (Adult Education) at Massey University. The proposed research study is designed to illuminate the assessment practices of tertiary Outdoor Education teachers.

It will involve you participating voluntarily in an interview/focus group to explore your assessment practices. You are invited to participate on the basis of your extensive experience in Outdoor Education and that you have been teaching at a tertiary institute for a number of years.

The results of this project may be published. Your responses will be treated confidentially by the researcher. Your name or identity will not be used or revealed in any written work. You can withdraw from the project or decline to answer any questions. You can gain access to the findings once the research is concluded.

I will be pleased to discuss any concerns you have about participation in the project.

My supervisor is: Gloria Slater
Tel 0800 MASSEY or 06 356 9099
Email G. R.Slater@massey.ac.nz

Special Note:
This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher named above is responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher, please contact their supervisor or Professor Sylvia Rumball, Assistant to the Vice-Chancellor (Ethics & Equity), telephone 06 350 5249, email humanethics@massey.ac.nz.

Appendix 5: Research Participant Consent Form

Research Question
What are the assessment practices of teachers of Outdoor Education qualifications in higher education programmes throughout New Zealand?
Sub-questions
1. What are the assumptions underpinning teacher assessment practices?
2. What assessment processes do they use and why do they use these?
3. What values are assessors communicating through their assessment practices?

Name of Researcher: Steve Chapman

Contact:

Information
I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I have the right to ask further questions at any time.

I understand I have the right to withdraw from the study at any time and to decline to answer particular questions.

I agree to provide information to the researcher on the understanding that my name will not be used without my permission.

This information will be used only for this research and publications arising from the research project.

I agree/do not agree to the interview being audio-taped. I also understand that I have the right to ask for the audiotape to be turned off at any time during the interview.

All tapes, computer discs, and paper copies of this information will be kept in a secure place while it is used for the purposes of transcription. All electronic copies of transcription will be erased once the project is completed. Transcription will occur by the researcher and/or by a trusted person who agrees to keep the information confidential. The transcriptions maybe viewed my supervisor.

You have the right to obtain a copy of the research findings.

Signed

Name.....

Date.....

Appendix 6: Massey Ethics Approval

Massey University

OFFICE OF THE ASSISTANT TO THE VICE-CHANCELLOR (RESEARCH ETHICS)
Private Bag 11 222
Palmerston North 4442
New Zealand
T 64 6 350 5573
64 6 350 5575
F 64 6 350 5622
humanethics@massey.ac.nz
animalethics@massey.ac.nz
gtc@massey.ac.nz
www.massey.ac.nz

31 July 2009

Steve Chapman
[redacted]

Dear Steve

Re: Illuminating the Assessment Practices of Outdoor Education Teachers in Higher Education

Thank you for your Low Risk Notification which was received on 30 July 2009.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

Please ensure that the following statement is included in all information provided to participants during recruitment (eg, information sheet, preamble to questionnaire, etc):

> *"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.*
>
> *If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor Sylvia Rumball, Assistant to the Vice-Chancellor (Research Ethics), telephone 06 350 5249, e-mail humanethics@massey.ac.nz".*

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

Sylvia Rumball

Sylvia V Rumball (Professor)
Chair, Human Ethics Chairs' Committee and
Assistant to the Vice-Chancellor (Research Ethics)

cc Ms Gloria Slater
School of Educational Studies
PN900

Dr Linda Leach
School of Educational Studies
PN900

Prof Howard Lee, HoS
School of Educational Studies
PN900

Ms Roseanne MacGillivray
Graduate School of Education
PN900