PRIFYSGOL
BANGOR
UNIVERSITY

Developing excellence in outdoor provision: Enhancing training pathways for outdoor qualifications

Thesis submitted to Bangor University in fulfilment of the requirements for the Degree of Doctor of Philosophy at the School of Sport, Health, and Exercise Sciences, Bangor University.

Will A. S. Hardy

2020-08-24

Declaration and Consent

I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

Funding

This PhD was funded by a Knowledge Economy Skills Scholarship (KESS 2), which was part funded by Mountain Training UKI and by the Welsh Government's European Social Fund.

Mae'r Ysgoloriaeth Sgiliau Economi Gwybodaeth (KESS 2) yn fentersgiliau lefel uwch Cymru gyfana arweinir gan Brifysgol Bangor ar ran y sector AU yng Nghymru. Fe'i cyllidir yn rhannol gan raglen cydgyfeirio Cronfa Gymdeithasol Ewropeaidd (ESF)ar gyfer Gorllewin Cymru a'r Cymoedd.

Knowledge Economy Skills Scholarships (KESS 2) is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government's European Social Fund (ESF) convergence programme for West Wales and the Valleys.

Open Science

https://github.com/w-hardy/thesis

Acknowledgements

Some words about some people.

Preface

Thesis Format

This thesis does not follow the traditional format; a substantial body of work is reported in appendices and empirical chapters have been written in a way that aids their preparation for publication, the main reasons for this are twofold. Firstly, a substantial amount of work was required to develop the data collection tool needed for the work reported in Chapter 3. This work was foundational in the PhD process, both developing the researcher and enabling five of the main studies, however including it in the main body of the thesis would distract and possibly detract from the "story" presented in the thesis. Therefore, this work is reported in an appendix so that the reader can engage with it having read the main body of the thesis, which reports the research focused on the research question. Secondly, this thesis aims to satisfy the dual objective of writing a thesis and preparing the research for publication. A consequence of writing this thesis with publication in mind is the self-referential terminology used in the empirical chapters. Accordingly, empirical chapters are written in first-person plural, consistent with conventions in reporting co-authored research, and the remaining chapters are written in first-person singular.

Terminology

Gender

It is important that we provide a note on terminology used in this thesis. Historically, and somewhat incorrectly, the terms *sex* and *gender* have been used somewhat interchangeably. Current guidelines from the American Psychological Association 2020 define gender as, "the attitudes, feelings, and behaviors that a given culture associates with a person's biological sex" (American Psychological Association, 2012, p 12); sex as, "biological sex assignment;" and *gender identity* as "a component of gender that describes a person's psychological sense of their gender" (American Psychological Association, 2020, Section 5.5).

Considering the definitions of sex and gender presented above, it would be more appropriate to use terms such as "man" and "woman" than "female" and "male" when discussing gender differences. However, in the studies reported in this chapter, participants' gender was retrieved from Mountain Training's Candidate Management System (CMS), where it is stored as female, gender neutral, and male. To avoid making presumptions about participants gender identities we have used the terms female and male throughout this chapter. No data were collected from gender neutral candidates, which may be unsurprising given that 0.03% of all Mountain Training Candidates report being gender neutral (Mountain Training UK, 2019). Further, following the same principal, when discussing previous research, we have used terminology consistent with that of the original authors.

Mountain Training

Mountain Training is not one single entity, it is in fact the outward facing name for a group of organisations in the United Kingdom and Ireland. Mountain Training England, Cymru, Scotland, and the Mountain Training Board of Ireland are responsible for

administering the Mountain Leader qualification in their respective countries. Mountain Training United Kingdom and Ireland are responsible for the generic training pathway for all qualifications. This project was funded through the KESS 2 programme with Mountain Training United Kingdom and Ireland as the partner organisation. However, as this project has stakeholders within each Mountain Training organisation, we simply refer to Mountain Training as a single entity throughout the thesis.

Contents

List of Tables

List of Figures

Thesis Abstract

Chapter 1

General Introduction

1.1 Mountain Training

Mountain Training is responsible for training instructors for walking, climbing, and mountaineering in the UK and Ireland. Its qualifications all follow a similar pathway to qualification, which was originally created in 1964 for the Mountain Leadership Certificate (what is now the Mountain Leader qualification) and has not changed much since then. Candidates must first gain some *prerequisite experience and register for the qualification*, then they *complete a training course*, following that they are required to gain *further experience to consolidate skills*, and finally they need to *successfully complete an assessment course*, following which they will be awarded the relevant qualification.

In 2018 there were 3,228 qualifications awarded to candidates, which suggests that this pathway is successful to some degree, as each year a large number of candidates are making it from registration to qualification. However, there is a drop off in the number of candidates at each step in the pathway for all qualifications (i.e., registration to training, training to assessment, and passing an assessment). This drop off can be seen in Figure 1.1.

It is estimated that Mountain Training qualification holders impact over 1.5 million people each year (Mountain Training England (2019)) and that number is likely to increase in the coming years as the "The overall growth in numbers [of active adults] continue [*sic*] to be driven by strong upward trends in walking and adventure sports (a category which includes hill and mountain walking, climbing and orienteering)" (Sport

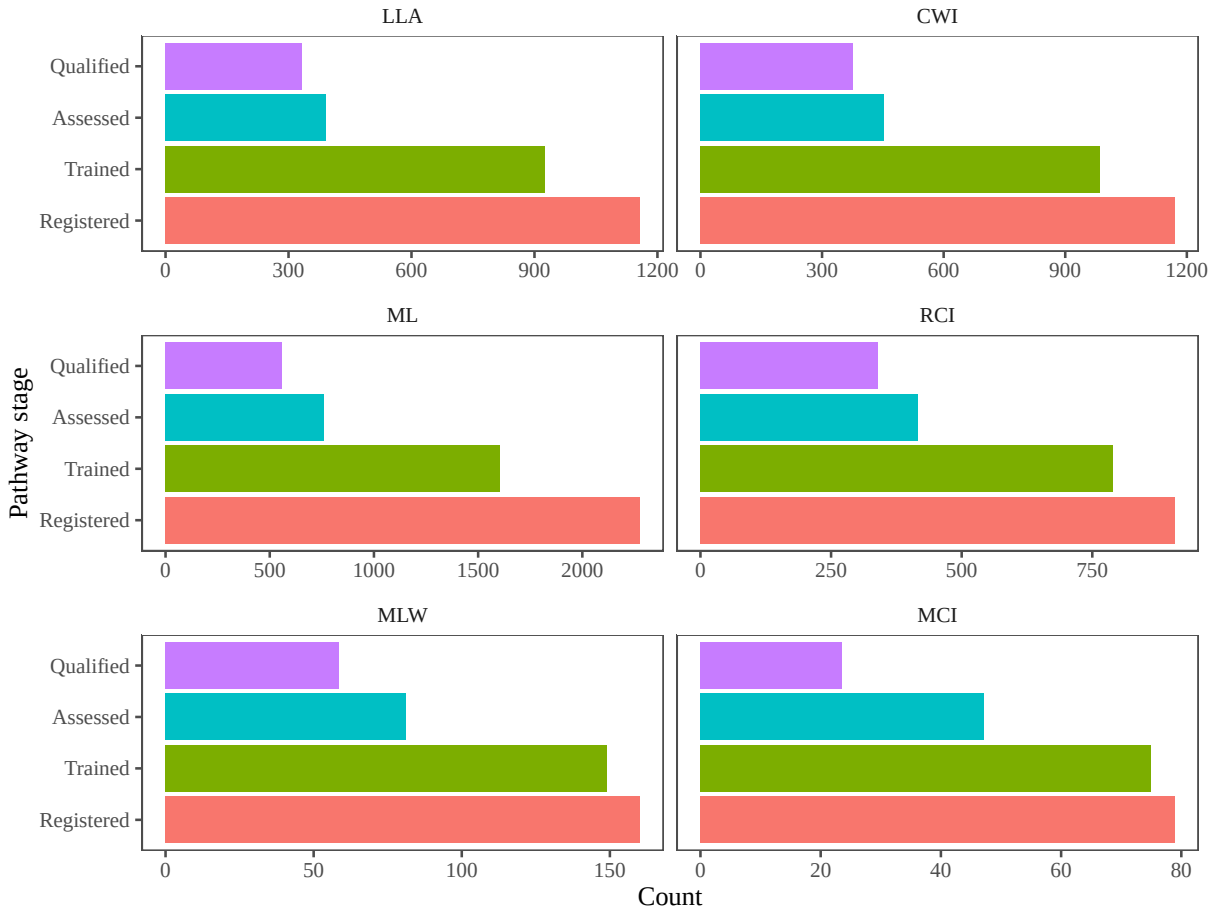

Figure 1.1: Average number of candidates at each pathway stage 2009-2018. LLA = Lowland Leader Award, CWI = Climbing Wall Instructor, ML = Mountain Leader - Summer, RCI = Rock Climbing Instructor, MLW = Mountain Leader - Winter, MCI = Mountaineering and Climbing Instructor.

England, 2020, p 14). Therefore, it is very important to Mountain Training that they understand why people do not complete their qualifications.

To begin to understand the factors influencing the completion of Mountain Training qualifications, this thesis focuses on the Mountain Leader qualification and does so for four main reasons: (a) it is the largest qualification as measured by number of candidates; (b) it has one of the largest drop-off in candidates progressing from training to assessment, the drop-off at this point is of particular interest as candidates have engaged with the Mountain Training delivery system; (c) it is the highest entry level qualification; and (d) it is the oldest qualification and has had few major changes to it recently.

1.2 The Mountain Leader Qualification

The Mountain Leader qualification is for "people who want to lead groups in the mountains, hills and moorlands of the UK and Ireland" (Mountain Training UK, 2015a, p 5). Whilst the Mountain Leader qualification is UK based and aimed at those who wish to lead others in the UK, training programmes in other countries are based on the success of the Mountain Leader qualification (D'Alpinisme, 2015). To qualify as a Mountain Leader, candidates must: (a) *register for the qualification* and gain a minimum of 20 Quality Mountain Days (QMDs),[1] (b) complete a six day *training course*, (c) gain a minimum of 20 additional QMDs as *further experience to consolidate skills*, and (d) successfully complete a five day *assessment course*. Therefore, to become a Mountain Leader, one must spend a minimum of 51 days in the mountains. Most successful candidates will have more experience than this, whether that is additional QMDs, experience of mountain walking that does not meet the QMD criteria, or other mountaineering experience. Becoming a Mountain Leader requires candidates to commit a significant amount of time and money in order for them to qualify as Mountain Leaders.

Between 2009 and 2018 an average of 2,278 candidates registered for the Mountain Leader qualification each year, but only 559 qualified a year. When looking

[1]There is not a simple definition for a QMD, however it is expected that a QMD will "make a positive contribution towards a person's development and maturity as an all round mountaineer" (Mountain Training, 2019)

more closely at the numbers of candidates who did qualify it became clear that there are two main components to qualifying: (a) getting to an assessment and (b) passing an assessment. Interestingly, most candidates who got to an assessment passed their first assessment (Figure 1.2), but most candidates did not get to an assessment (Figure 1.3). It is also noteworthy that, as shown in Figure 1.3, becoming a Mountain Leader is not a quick process (period between training and assessment, $M = 1.57$ years, $SD = 1.45$).

To examine this difference for the Mountain Leader qualification in more detail we carried out a *survival analysis*, where rather than looking at summary statistics averaged over a number of years, we look at the probability of an individual candidate having been assessed over time following their training course. As can be seen in Figure 1.3 at any given point in time, fewer female candidates get to an assessment than male candidates. The percentage likelihood of a candidate having been assessed five years following their training course is ~32% and ~40% respectively for female and male candidates, after this point the rate of candidates being assessed decreases for both genders Over half of candidates who did reach assessment did that within 18 months of their training courses, but it is not unusual to take longer, and some candidates do go on to be assessed over five years after their training course.

We also examined the pass rates for the Mountain Leader qualification. The pass rate is increasing over time and there have been changes to sex differences in the pass rates over the last 10 years (Figure 1.2). When looking at pass rates for the last 10 years, women were less likely to pass their first assessment, but the pass rate was increasing faster for them than it was for men. However, when looking at data from the last five years, neither the effect of sex on the pass rate or rate of change of the pass rate are statistically significant.

Mountain Training would like to understand why more people do not complete the Mountain Leader qualification and identify if there are any changes to the pathway that they can make in order to better support their candidates. It is unlikely that there is a single factor that would be a “silver bullet” to improve completion rates. Instead there are likely a myriad of factors which influence completion at various stages of the pathway. Some of these factors will be generic to all candidates, whilst some may be specific to individual (groups of) candidates.

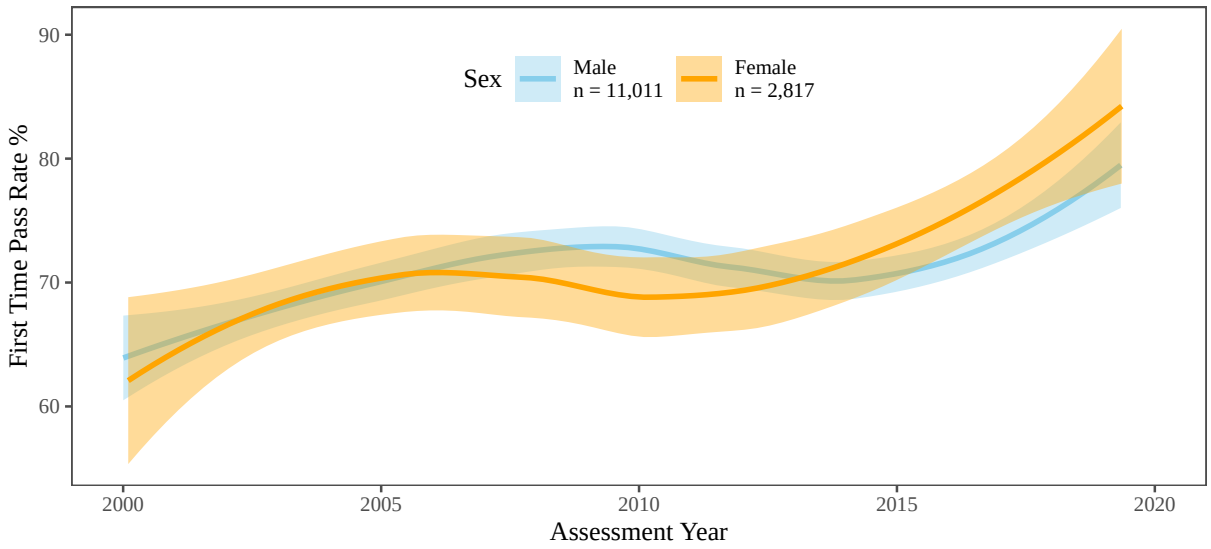

Figure 1.2: Pass rates for female and male candidates assessed since 2000 ($N = 13,828$).

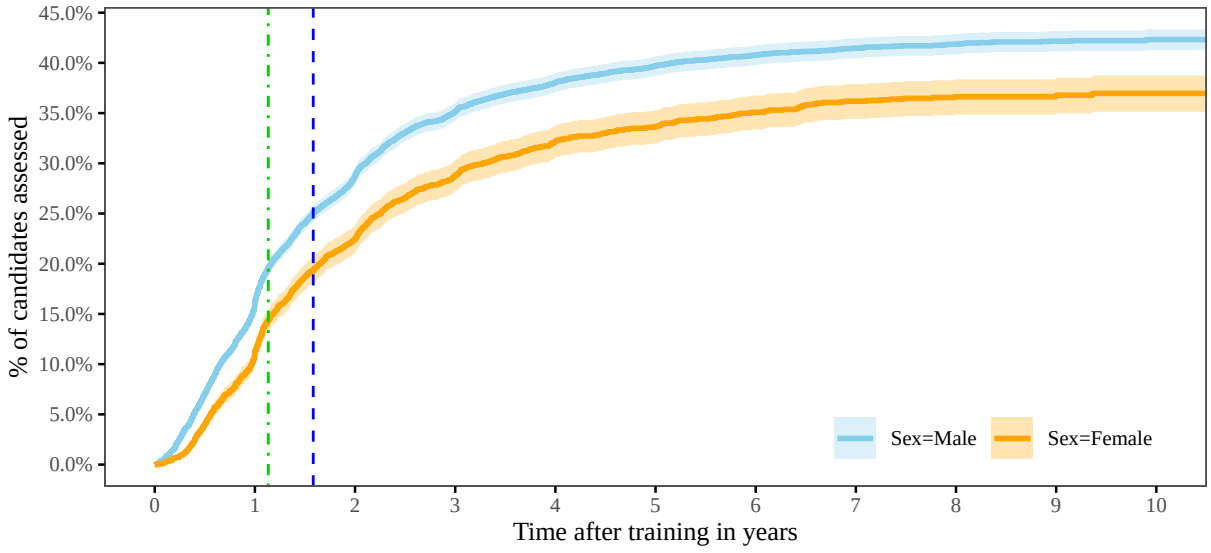

Figure 1.3: Survival rates for female and male candidates post-training. Candidates trained 2009-2019 ($N = 15,635$). Green dotted and dashed line at median time to assessment and blue dashed line at mean time to assessment.

1.3 Thesis Structure

- Chapter 1 has outlined the rational for this thesis and it introduces potentially relevant theory below. and develops research questions to be explored throughout the thesis.
- Chapter 2 presents a qualitative inquiry that aimed to explore organisational managers' understanding of factors that influence completion of the Mountain Leader qualification.
- Chapter 3 comprises three studies, each of which use state of the art pattern recognition techniques to identify key discriminatory features that provide insight into the factors influencing different stages of the qualification pathway.
- Chapter 4 explores the main and interactive effects of gender and experience on self-efficacy. This involved two studies, the first was the development and validation of a measure that could be used to measure Mountain Leader related self-efficacy and the second was a study which used this tool to test the relationships of gender and experience with self-efficacy.

1.4 Relevant Theory

It is beyond the scope of this thesis to provide a comprehensive review of all relevant literature, however, below are brief descriptions of theories which provide useful frameworks for discussing the results of each chapter. Whilst these theories are useful for understanding the relationships between various factors, this thesis is not a test of any specific theory.

1.4.1 Self-Determination Theory

Self-determination theory (Deci and Ryan, 1985a, 2000; Ryan and Deci, 2017) is a theory of human behaviour with six mini-theories that sit within it: *cognitive evaluation theory* (Deci, 1975; Deci and Ryan, 1980), *organismic integration theory* (Deci and Ryan, 1985a; Ryan and Connell, 1989), *causality orientations theory* (Deci and Ryan, 1985b), *basic psychological needs theory* (Ryan and Deci, 2000b), *goal contents theory* (Kasser and Ryan, 1996; Niemiec et al., 2009), and *relationship motivation theory* (Deci and Ryan, 2014; Ryan and Deci, 2017). Cognitive evaluation theory, organismic integration

theory, and basic psychological needs theory are particularly relevant to the present study and are described below.

1.4.1.1 Cognitive Evaluation Theory.

Cognitive evaluation theory is concerned with the relationships between social factors (e.g., rewards, motivational strategies), which may be autonomy supportive, controlling or amotivating, and intrinsic motivation and interest (Deci, 1975; Deci and Ryan, 1980). Cognitive evaluation theory suggests that competence and autonomy will foster intrinsic motivation. Social factors that are perceived as *controlling* will lead to an external *perceived locus of causality*, frustrating individual's need for autonomy, thus undermining intrinsic motivation; where as social factors that are perceived as *informational* support perceived competence, thus enhancing intrinsic motivation (Ryan and Deci, 2019).

1.4.1.2 Organismic Integration Theory.

Organismic integration theory is concerned with extrinsic motivation and activities where the outcome is separable from the behaviour (Ryan et al., 1985).
Self-determination theory traditionally suggests that every behaviour can be placed on a continuum, the relative autonomy continuum, with autonomous or self-determined motives on one side and controlled or non-self-determined motives on the other, and that people will therefore vary in both levels of and quality of motivation (Deci and Ryan, 1985a; Sheldon and Prentice, 2019).[2]

Organismic integration theory suggests that there are four types of extrinsic motive—despite suggesting that they form a continuum. Two of these are considered controlled forms of extrinsic motive and the other two are considered autonomous forms of extrinsic motive. The most controlled form of extrinsic motivation proposed is *external regulation*, where an individual is motivated by external pressures or reward and whist it can be a powerful form of motivations, it is not typically thought of as long lasting (Ryan and Deci, 2000a). The next form of extrinsic motive is *introjected regulation*, where behaviour is motivated by internally controlling pressures (e.g., ego-involvement, contingent self-worth). This form of motivation may also be powerful,

[2]There is some evidence that the various types of motivation should not be placed on a continuum but may in fact be better considered as contiguous (Chemolli and Gagné, 2014). That notwithstanding, it is still considered that the different types of motivation may have different outcomes.

but may be weakened in the face of setbacks (Deci and Ryan, 1995; Ryan and Deci, 2019).

Considering the more autonomous forms of extrinsic motivation, *identified regulation* is when an individual consciously values their engagement in the activity and accepts the behaviour as personally important. Finally, the most autonomous form of extrinsic motivation proposed is *integrated regulation*, when the behaviour is congruent with personally endorsed values. Integrated regulation shares features with intrinsic motivation, however, are considered extrinsic as the outcome is separable from the behaviour (Ryan and Deci, 2002). Self-determination theory would normally suggest that autonomous forms of motivation are more sustainable and better predictors of performance and goal persistence than controlled forms of motivation are (Hagger and Chatzisarantis, 2015; Pelletier et al., 2001; Ryan and Deci, 2019).

Organismic integration theory suggests that social factors that support the basic psychological needs foster the development of more autonomous forms of extrinsic motivation, however it does not suggest that the relative autonomy continuum is a developmental continuum (Deci and Ryan, 1991; Ryan, 1995; Ryan and Deci, 2002). Autonomy support, the provision of structure, and involvement have been identified as need supportive behaviours, which support the development of autonomous regulation (Ryan et al., 2016; Markland and Tobin, 2010).

1.4.1.3 Basic Psychological Needs Theory.

Basic psychological needs theory is central to self-determination theory and each of these mini-theories (Ryan and Deci, 2002). Basic psychological needs theory posits three basic psychological needs: *autonomy*, feelings of volition, choice, and internal control; *competence*, the feeling of mastery and effectiveness; and *relatedness*, feeling connected and involved with others. Self-determination theorists suggest that the satisfaction of all three of these needs supports well-being and "high-quality" motivation (Ryan and Deci, 2019) and is essential for optimal-functioning, good mental health, and well-being (Chen et al., 2015; Deci and Ryan, 2000).

1.4.1.4 Hierarchical Models of Motivation.

Researchers have suggested that motivation operates in a hierarchical fashion (Ingledew et al., 2009; Vallerand, 1997; Vallerand and Ratelle, 2002). These researchers suggest that motivation is a complex construct and differs in types and exists on at least three different levels: the global/dispositional, why an individual generally engages in activities; contextual/participatory, the contents of motives within a particular domain or what an individual is trying to achieve or avoid; and situational/regulatory motives, the perceived locus of causality of the behavioural goals—where the motive sits on the relative autonomy continuum (Deci and Ryan, 2000; Ingledew et al., 2009; Vallerand, 1997; Vallerand and Ratelle, 2002). Whilst at the global level, motivation is considered to be an individual difference and therefore results in general consequences, at the participatory level, the context will influence motivation and may be manipulated more easily (Vallerand and Ratelle, 2002).

The results from the qualitative study suggested that motives for becoming a Mountain Leader existed on at least two of these levels, the participatory and regulatory. Therefore, in this study we measured two levels of motivation proposed by self-determination theorists: participatory motives and regulatory motives the perceived loci of causality of behavioural goals. In the qualitative study, all participants suggested that candidates (both female and male) with extrinsic participatory motives were more likely to be assessed than those who had intrinsic participatory motives.

1.4.2 Self-Discrepancy Theory.

Self-discrepancy theory (Higgins, 1987) considers three types of self: the actual-self is the way in which an individual perceives themselves, the ideal-self represents their hopes and aspirations, and the ought-self represents their obligations and responsibilities. The theory posits that individuals will be motivated to reduce discrepancies between their actual-self and their different self-guides (ideal and ought) and that different discrepancies will result in specific types of emotions: actual/ideal discrepancies will result in dejection-related emotions (e.g., sadness, disappointment, dissatisfaction) and actual/ought discrepancies will results in agitation-related emotions (e.g., guilt, worry, tension).

1.4.3 Self-Efficacy Theory.

Self-efficacy refers to an individual's confidence in their ability to carry out a specific task at a given time. Self-efficacy theory suggests that if an individual possesses the necessary skills and is sufficiently motivated, then their level of self-efficacy will be the primary determinant of their performance, how much effort they will put in, and how long they will persist—particularly in the face of adversity (Bandura, 1977, 1982, 1997). Bandura (1977) suggested that an individual's efficacy beliefs could vary in level, strength, and generality.

Meta-analyses have found evidence supporting the relationships suggested by Bandura between self-efficacy and performance and persistence across several domains, for example in sports (Moritz et al., 2000), in work contexts (Stajkovic and Luthans, 1998), and academic environments (Multon et al., 1991). Several experimental lab studies have that manipulated participants' level of self-efficacy found that higher levels of self-efficacy are related to increased task persistence (Hutchinson et al., 2008; Tenenbaum et al., 2001; Weinberg et al., 1979, 1980, 1981). However, a meta-analysis of within-person studies across various domains found evidence that supported the hypothesised relationship between previous experiences and self-efficacy, but suggested that the relationships between self-efficacy and performance in between-person studies may be a result of those with higher levels of self-efficacy having been successful in the past (Sitzmann and Yeo, 2013).

Bandura (1982) suggested that an individual's perception of self-efficacy will be based on four main sources of information:[3] previous performance accomplishments, vicarious experience/modelling, social/verbal persuasion, and physiological/emotional states.

Performance accomplishments will be the strongest source of information as it is based on one's mastery experiences; successful experiences will increase perceived self-efficacy and repeated failures will decrease it, especially if failure occurs early and are not attributed to a lack of effort or external circumstances. Previous efficacy beliefs will influence how an individual integrates an experience; experiences that are congruent with previous efficacy beliefs will be integrated more readily than those which are

[3]Other researchers (e.g., Maddux, 1995) have added additional sources, however, most research considers the four sources proposed by Bandura.

incongruent (Bandura, 1997; Cervone and Palmer, 1990; Chase, 2001). Bandura (1997) also suggested that weak efficacy beliefs will be more easily reduced by disconfirming events. Research on self-efficacy for future physical activity—across a range of age groups—supports the hypothesis that that self-efficacy beliefs for future participation are influenced by previous performance (e.g., Dawson and Brawley, 2000; McAuley et al., 2006; Miller and Mulligan, 2002).

In addition to gleaning information from their own experiences, individual's self-efficacy will be informed by *vicarious experiences* also known as *modelling*; seeing other people, similar to oneself, successfully carry out a task can increase an individual's level of self-efficacy (Bandura, 1982; Lirgg and Feltz, 1991), and seeing them fail may lower their efficacy expectations (Brown and Inouye, 1978). Vicarious experiences provide less information about efficacy, but are still important (McCullagh and Weiss, 2001) and may be an especially important source of information to those who lack experience and therefore, their own mastery experience (Weinberg and Gould, 2014). Bandura (1986) suggested that the greater the similarity between the individual and the model, the greater the potential for the model to influence the individual's efficacy beliefs. However, empirical research had yielded equivocal findings when testing this hypothesis (see Samson and Solmon, 2011).

Social/verbal persuasion are techniques used by the self and others, usually with the aim of improving confidence and can lead to improved performance assuming that the individual's resultant efficacy beliefs are realistic (Bandura, 1982). Social/verbal persuasion can come in different forms including, feedback, social support, goal setting, and self-talk (Samson and Solmon, 2011).

Physiological/emotional states both influence efficacy beliefs (Bandura, 1977, 1982). If an individual associates physiological arousal with poor performance their efficacy beliefs will be reduced, but if there is a positive association with that physiological arousal then their self-efficacy will be enhanced (Hauck et al., 2008; Jones, 1995; Jones and Swain, 1992). There is also evidence that positive emotional states are more likely to lead to positive efficacy beliefs than negative emotional states are (Maddux, 1995; Martin, 2002).

1.4.4 Theory of Planned Behaviour

The theory of planned behaviour (Ajzen, 1991; Ajzen and Madden, 1986) is a model for predicting behaviour and builds on the theory of reasoned action (Ajzen and Fishbein, 1980). The theory of reasoned action suggests that intentions are the best predictor of future behaviour. It suggests than an individual's intentions are the result of their attitudes towards the behaviours and the *subjective norm* surrounding that behaviour. Subjective norms are formed based on an individual's belief about other's views and the motivation of the individual to comply with those views. The theory of planned behaviour extends the theory of reasoned action by including *perceived behavioural control* as an influence on both the intention and the behaviour. An individual's perceived behavioural control is extent to which they feel that they have the resources and opportunities to control the behaviour.

Meta-analyses have found evidence that the constructs that make up the theory of planned behaviour can predict future behaviour (Armitage and Conner, 2001; Hagger et al., 2002). However, the theory of planned behaviour is not without its critics, some of whom suggest that the theory is too limited in the constructs that it includes (cf. Sniehotta et al., 2014).

1.4.5 Integrated Models

All of the theories introduced above have substantial bodies of research behind them, however several researchers have sought to better explain phenomena of interest by integrating more than one of these theories into new models (e.g., Fishbein, 2000; Hagger and Chatzisarantis, 2009, 2014; Hamilton et al., 2017). The *integrated behaviour change model* (Hagger and Chatzisarantis, 2009, 2014; Hamilton et al., 2017) suggests that motivation will cause attitudes, subjective norms, and perceived behavioural control in the theory of planned behaviour. In a model including various individual difference variables, Fishbein (2000) replaced the perceived behavioural control construct of the theory of planned behaviour with self-efficacy (n.b., Ajzen, 1991, suggested that perceived behavioural control is similar to the construct of *self-efficacy*) and included skills and environmental constraints as addition caused of behaviour.

Chapter 2

Study 1: Qualitative

2.1 Introduction

A review of relevant literature identified 52 factors that may influence the completion of the Mountain Leader qualification. This was deemed to be too many topics to include in the interviews, therefore we carried out an expectancy value ranking process to identify the factors that the research team felt would be most important.

The ease of collecting the data for each factor was rated on a three-point scale, from *difficult* (1) to *easy* (3). The expected utility of each factor was also rated on a three-point scale, from *low* (1) to *high* (3). WH rated the factors initially, then discussed the ratings with RR and discussed any disagreements till a consensus was found. We then calculated the product of these two ratings to create a single score for the expected likelihood of collecting useful data for each variable (see Table 2.1). This score was then used to rank variables, with factors scoring less than six being removed. The remaining 44 factors were then used to develop a guide for the interviews.

It is important to note that we considered utility rather than perceived importance in this exercise. For instance, mental toughness, which was defined by Bell et al. (2013) as "the ability to achieve personal goals in the face of pressure from a wide range of different stressors" was not included in the interview guide as we felt that it would be somewhat tautological to ask a question along the lines of "are people who are good at completing things good at completing things?" The factors retained, were split into five sections: candidate background, candidate career history and social influence, candidate personal characteristics, candidate experience and ability, and candidate

support.

2.1.1 Research Questions

The broad questions considered within each section of the interview are described below. In addition, the full interview guide is presented in Appendix A. To aid readability of this chapter and understanding of the results, we introduce relevant literature at the start of each theme within the results before presenting the data from the interviews.

2.1.1.1 Candidate Background.

Does a candidate's socio-economic background influence their progression through the Mountain Leader qualification? How does a candidates age influence their progression? Are there any professions that help or hinder candidates progression?

2.1.1.2 Candidate Career History and Social Influence.

Why do people want to become Mountain Leaders? Do these reasons influence their progression? How long do candidates think it will take to become Mountain Leaders? Do candidates see becoming a Mountain Leader as a stand alone profession or do they intend to use it along side another job? How do candidates who want to continue on to higher-level qualifications differ to those who only want to be Mountain Leaders? Do candidates have role models? What influence do role models have on candidates?

2.1.1.3 Candidate Personal Characteristics.

What attitudes do candidates have towards the Mountain Leader qualification? How confident are candidates that they can become Mountain Leaders? What increases/decreases this confidence? Do candidates want to be as good a they can be or just good enough? What sort of disconfirmatory experiences do candidates have? How do these experiences affect different candidates? How do candidates cope with setbacks?

2.1.1.4 Candidate Experience and Ability.

How does prior experience influence performance at assessment? What types of experience help/hinder candidates becoming Mountain Leaders? What causes candidates with lots of experience to perform poorly at assessment? What causes

Table 2.1: Expectancy value exercise.

Factor	Ease of collection	Expected utility	Expectancy value
Candidate background			
Age at registration/life history	3	3	9
Location	3	3	9
Age	3	2	6
Sex	3	2	6
Social media profile	2	3	6
Level of education	2	2	4
Socioeconomic status	1	2	2
Candidate career history and social influence			
Intention/goal expectations	3	3	9
Opportunities for relevant work	3	3	9
Subjective norms/social influence	3	3	9
Relevant sports media influence	3	3	9
Role model	3	3	9
Engagement with “mainstream sports”	3	2	6
Other qualifications	3	2	6
Profession	2	3	6
Critical developmental experiences	2	3	6
Opportunities to participate	2	2	4
Enjoys exercise	2	2	4
Number of career changes	2	2	4
Attraction to everyday life vs. outdoor life	2	1	2
Mid-career turning point - injury	2	1	2
Candidate personal characteristics			
Achievement orientation	3	2	6
Conscientiousness	3	2	6
Grit	2	3	6
Attitudes and outcome expectations	2	3	6
Perfectionism	2	3	6
Optimism	2	3	6
Motivation - SDT	2	3	6
Resilience	2	2	4
Need to succeed	2	2	4
Obsessiveness	2	2	4
Mental toughness	3	1	3
Locus of control	1	3	3
Selfishness	2	1	2
Family values	1	2	2
Ruthlessness	1	1	1
Candidate experience and ability			
Training staff/centre	3	3	9
Personal experience	3	3	9
Personal competency	2	3	6
Perceived competence	2	3	6
Self-efficacy	2	3	6
Disconfirmatory experiences	2	3	6
Family activity	3	1	3
Candidate support			
Mentoring	3	3	9
Association membership	3	2	6
Instructor support	2	3	6
Family/significant other support	2	3	6
Peer-support	2	3	6
Emotional support	2	3	6
Esteem support	2	3	6
Informational support	2	3	6
Tangible support	2	3	6

candidates with little experience to perform well at assessment? How do candidates views gaining experience? How does a candidates level of experience affect them on a training course?

2.1.1.5 Candidate Support.

What do course staff think about the Mountain Leader qualification? How do course debriefs influence candidates? Do course staff help candidates plan their progression? What are course staff's coaching and leadership behaviours like? How prevalent is mentoring? What is good/bad about mentoring? Which types of support help/hinder candidates progression? Where do candidates get support from? Are candidates supported well enough? Which candidates need more/less support?

2.2 Methods

2.2.1 Philosophical Orientation

We adopted a relativist epistemology for this study, as we believe that different realities will exist for individuals on different "sides" of the Mountain Leader qualification (e.g., candidates and course staff; Sparkes and Smith, 2009). A relativist epistemology is incongruent with the *parallel perspective* (Sparkes, 1998), where a researcher would seek to establish validity in a similar way to that of a quantitative study. Instead, we present the reader with evidence of the credibility of this research, allowing them to judge the rigour of the study for themselves (Sparkes and Smith, 2009).

2.2.2 Participants

After gaining ethical approval from Bangor University's School of Sport, Health, and Exercise Sciences ethics committee, and individual informed consent, seven individuals participated in this study. Initial interviews carried out with four members of staff from Mountain Training including staff from: Mountain Training Cymru, England, and Scotland, the three main national training boards and Mountain Training United Kingdom and Ireland (three men and one woman; age, $M = 47.19$ years, $SD = 6.60$; number of Mountain Leader courses worked $M = 41.60$ courses, $SD = 29.86$, $range =$

13-78).

Having completed these interviews it became clear that it would be important to interview course staff who had a greater knowledge of candidates and their experiences than the Mountain Training staff had. Therefore, we interviewed two high volume course providers and a course director who had worked for eleven different providers over 14 years (two men and one woman; age, $M = 55.30$ years, $SD = 5.18$; number of Mountain Leader courses worked $M = 284.00$ courses, $SD = 214.68$, $range = 66\text{-}576$). We used a purposive sampling strategy in this study (Patton, 2002; Sparkes and Smith, 2014). The individuals who participated in the study were recruited based on their knowledge and experience of the Mountain Leader qualification from an organisational perspective, rather than their personal experience of becoming a Mountain Leader.

2.2.3 Semi-Structured Interviews

It is important to match the research method to the question being asked (Smith and Sparkes, 2016). For this study we chose to use semi-structured interviews as they can provide a rich, yet broad, understanding of a given phenomenon (Lincoln and Guba, 1985). To facilitate discussion and ensure that all participants were asked broadly similar questions, covering the same topics, we developed an interview guide which included questions that covered the 44 factors deemed as the most important following the expectancy value ranking process described above. The interview guide was split into five sections: (a) candidate background, (b) candidate career history and social influence, (c) candidate personal characteristics, (d) candidate experience and ability, and (e) candidate support.

Each section began with questions designed to help participants focus their attention on the topic of interest (e.g., “Could you start by describing a typical group of six candidates on a Mountain Leader training course to me?”). Participants were then asked more specific questions, which related to the factors identified in the expectancy value exercise (e.g., “Are there are any professions that significantly influence, positively or negatively, completion of the Mountain Leader award?”). These questions were followed up with elaboration probes (e.g., “What do you think it is about these professions that makes a difference here?”) to improve the clarity and detail of the data (Patton, 2002).

Each section ended with two final questions. Firstly, a question asking if there was anything else that participants thought was relevant to the completion of the Mountain Leader qualification but had not been discussed (e.g., "Is there anything about candidates' backgrounds that you think is important but we haven't spoken about"). Secondly, a question that asked if they felt that any of the topics discussed in that section were more salient than the others (e.g., "We have spoken about a number of different factors relating to candidate background. Do you think that there are any factors relating to candidate background that are generally more important with regards to completion of the Mountain Leader award?").

Eight pilot interviews were conducted with Mountain Leader course staff and a Mountain Training Officer to familiarise the interviewer with the interview guide, to identify any factors not included in the interview guide that may be important, and to ensure that participants were able to provide sufficiently detailed answers to the questions. Minor changes were made to the final interview guide following each pilot interview, the final interview guide can be found in Appendix A.

WH was instructed in qualitative research methods by the RR and LH and additional knowledge was gleaned from recent literature on qualitative research methods (e.g., Smith and McGannon, 2018). The research team all had mountaineering experience and an understanding of the Mountain Leader qualification at the time the interviews were conducted. This experience allowed us to be more sensitive to the specific experiences and language of the participants, reducing the likelihood of introducing bias thorough insinuation and assumption (Denzin and Lincoln, 2005).

WH, who had 10 years of outdoor experience across the world, conducted all the interviews. Most of this experience was gained in a recreational context rather than a professional one, this is seen as a strength because the analysis was less likely to be influenced by personal experiences of Mountain Training qualifications, ensuring that it is the participants experiences that are presented. In addition to this, RR is a senior lecturer in sports psychology with over 20 years of outdoor experience and LH is a professor in sports psychology and has over 50 years of outdoor experience and is an IFMGA Mountain Guide. The experiences of the research team meant that good rapports could be established with participants and that the subtleties of the phenomena of interest could be fully understood.

2.2.4 Procedure

All interviews were carried out face-to-face in a location chosen by the participants (e.g., home, office space or a café). Given the exploratory nature of the interviews and range of factors included in the interview schedule the interviews were carried out over two to five sessions to avoid fatigue of the participants and interviewer (duration, $M = 316.25$ min, $SD = 54.85$). The interviewer made notes during the interview and the interviews were digitally recorded. The recordings were transcribed clean verbatim by UK Transcription yielding 314,927 words, or 1,329 transcript pages. Before beginning the analysis, the first author listened to the recordings whilst checking the transcripts for errors.

2.2.5 Data Analysis

The transcripts were analysed using thematic analysis in a concurrent inductive ("bottom up"/data driven) and deductive ("top down"/theory driven) fashion (Braun and Clarke, 2006). Analysing the data using a concurrent inductive and deductive approach allowed us to consider the data in relation to existing theory, but also to create new themes from the data. The flexibility of this analytical approach was important to this study as we were trying to identify potentially important factors, some of which we may not have considered *a prioi*. Acknowledging the existence of relevant literature and relating the data to it whilst also considering new themes of interest allowed us to make best use of the rich data that were collected without sacrificing its complexity and nuance.

The analysis of the transcripts was carried out in a series of separate steps. First, WH read each transcript to familiarise himself with the data. Following this, he coded the transcripts using NVivo 11 Pro (QSR, 2017) into the five *a priori* deductive codes and a sixth code—"Other." This sixth code was then analysed in an inductive fashion to identify any themes not encompassed by the deductive codes. Once all first-order themes had been identified, sub-themes were identified within each theme. WH presented a summary of each theme, including raw quotes from the interviews, to RR and LH who acted as *critical friends* (cf. Sparkes and Smith, 2014; Smith and McGannon, 2018). Acting as critical friends, RR and LH offered critical feedback, the aim of which was not to reach a consensus, but to encourage reflexivity (Smith and McGannon, 2018).

2.3 Findings

When conducting the interviews, it quickly became apparent that becoming a Mountain Leader has at least two distinct steps: firstly, a candidate must *get to an assessment* and secondly, they must *pass an assessment*. If they fail to pass their first assessment then they have the opportunity to return for a *reassessment*, which they may or may not pass. Consequently, we present the findings under three main headings: (a) getting to assessment, (b) passing, and (c) reassessment. This is done as there are differences in the factors that participants felt were important at each step, but also to aid the readability of this Chapter.

2.3.1 Getting to Assessment

Factors that influenced the likelihood of a candidate being assessed could be considered under three main themes: confidence, motivation, and gaining experience. There were three additional themes that participants felt influenced whether candidates reached assessment (albeit to a lesser degree): candidate location, reengaging later in life, and redirection to lower qualifications.

2.3.1.1 Confidence.

The results in this section shows that candidates must be confident in their ability to pass a five-day Mountain Leader assessment before they will attend one and the threshold of confidence required is individual to them. There are several factors that influence both the level of confidence and the individualised thresholds that candidates must surpass. Level of confidence.

2.3.1.1.1 Level of Confidence. All seven participants said that candidates needed to feel confident before they would attend an assessment. Officer 1 said:

> [Candidates] have to put themselves on a little pedestal and go, "This is me, and this is how I'm trying to go through the scheme" …. That takes someone who's got a reasonable amount of confidence in themselves to do that. I can imagine some candidates not feeling comfortable in putting themselves in that position … and I think that they will be the ones less

likely to complete.

Officer 2 supported this when describing candidates on assessment courses by saying, “In their heads, they’re prepared for it.” suggesting that only well-prepared, and therefore confident, candidates attended assessment courses. Talking about candidates who did not attend assessments Provider 1 said, “They convince themselves they’re not ready, and then they won’t book on.” However, Officer 1 suggested that more than just experience was needed for candidates to feel confident when he talked about a candidate who, “Doesn’t have the confidence to do the assessment” despite them being a “Super keen hillwalker … who has done the training.”

This evidence shows that the candidates who have reached assessment were confident in their abilities and that some of those who have not reached assessment did not feel confident, it also suggests that their confidence was not always dependant on their abilities.

2.3.1.1.1.1 Individual Differences in Thresholds of Confidence. This section presents evidence that candidates have their own thresholds for confidence that they must surpass before they will attend an assessment together with factors that influence that threshold, thus moderating the relationship between level of confidence and the likelihood of booking an assessment.

Five participants suggested that younger candidates have lower thresholds for confidence and that older candidates were less likely to feel confident enough to attend an assessment. Officer 4 said, “younger folk can be less constrained by lack of confidence.” Provider 2 supported this, suggesting that if older candidates did not feel confident they are more likely to refrain from booking an assessment, “Some of the older guys and girls have come in already with 40 days but they still might not come back for a year or two because they’re sometimes not as confident.” Officer 1 and Officer 2 did not comment on how age may or may not relate to confidence and getting to assessment.

Five participants discussed the effect of gender on confidence and all said that females needed to be more confident than males to attend an assessment. When asked if there were many candidates who were ready for their assessment but did not feel ready, and so did not attend an assessment, Provider 1 said:

I think a lot of girls fall into that category. That they actually could do it,

> but it feels like such a big thing. They want every "i" dotted and every "t" crossed, and they want to be absolutely doubly sure that they can do it, and really, they could have done it earlier.

Five participants suggested that some candidates may not have attended an assessment because their perfectionistic traits led to them having a higher threshold for confidence, thus not feeling confident enough to attend and assessment. Provider 2 gave an example where high-levels perfectionistic strivings may have led to candidates not feeling ready for an assessment despite being ready, "For some reason or another, they've really held back … it could be that they're an absolute perfectionist and they just didn't want to turn up until they were totally happy."

Officer 3 suggested that female candidates had higher levels of perfectionistic concerns, thus were more likely to incorrectly feel that they were "below the standard", "A female might actually be overly cautious about exposing themselves, and potentially failing … through believing they're actually below the standard. Whereas they're probably higher than that." Officer 1 and Provider 3 did not discuss how perfectionism may or may not influence candidates' confidence threshold.

2.3.1.1.1.2 Understanding the Standard. Throughout the interviews all participants referenced "the standard" (i.e., the standard required to pass) and commented that it is often hard for candidates, and sometimes staff, to understand what it is. The five participants that discussed "the standard" and how it related to getting to assessment all suggested that a candidate's confidence level may not surpass the threshold needed to attend an assessment because they did not understand "the standard," thus making it hard to be confident. Provider 3 explained that this holds some candidates back from being assessed:

> They need that reassurance that … they're consolidating correctly, and actually they're performing at the standard … because they're not going to come forward unless they feel like that …. I think that's really hard [for candidates] to know where they're at in relation to the bar. We think it's clear … but candidates always ask, "How close do I have to be? You know, like, ten metres, a hundred metres. One contour line, two contour lines."

Three other participants made similar comments, and Officer 2 did not refer to understanding the standard.

2.3.1.1.1.3 Raising Candidates' Confidence Levels. Six participants discussed how support helped close the gap between candidates' confidence levels and their confidence thresholds by raising confidence levels rather than lowering confidence thresholds. When talking about candidates who lacked confidence Provider 3 said:

> They often need a lot more support, and with a bit of support they often shine as well: as soon as they realise that, "Actually, I am good enough and I can do it", then they're up and running, although it can be fragile, and it doesn't take much to knock it.

Talking about isolated candidates Officer 4 said, "I suppose the ones without a network … and those at the lower end of the confidence spectrum … are going to need help with upskilling or believing that they've got the skills in the first place."

2.3.1.1.2 Gender Differences in Robustness of Confidence. In addition to the gender differences in confidence thresholds discussed above, three participants spoke about gender differences in the robustness of candidates' confidence. When talking about the different influences of negative events on candidates, Officer 3 said, "Who can take it in their stride? Blokes, I suppose. Not because they perform well afterwards, they will probably be weaker. They are more Bolshie, I suppose." Officer 2 supported this:

> For some candidates, particularly men, those effects of that bit of negative feedback or that bad day they had on the hill, they try and brush off and just carry on … and get it right next time …. What you might find with many females is that's thrown a spanner in the works. It's made them doubt what they need to do, and now they need to readjust their consolidation plan.

2.3.1.2 Motivation.

Many motivation researchers have proposed that motivation operates at different levels (e.g., Vallerand, 1997; Vallerand and Blssonnette, 1992). In particular,

self-determination theorists have proposed three levels of motivation: dispositional motives (i.e., goals for life in general), participatory motives (i.e., what someone hopes to achieve or avoid by participating in a behaviour), and regulatory motives (i.e., the perceived loci of causality of behavioural goals) (e.g., Deci and Ryan, 2000; Ingledew et al., 2009). All seven participants gave examples of candidates with different participatory motives who also had different regulatory motives within those participatory motives. They suggested that both levels of motive influence candidates' likelihood of attending assessment.

2.3.1.2.1 Participatory Motives. All participants said that candidates with extrinsic participatory motives (i.e., to achieve an external goal), particularly allowing them to work, are more likely to complete than those with participatory intrinsic motives (i.e., doing something for its own sake). Provider 1 said "The ones where there's a driver, are more likely to [complete] …. If they're not doing it for work and they're using it in an informal thing, they are probably less likely to [complete]." Provider 1 went on to say, "People who want to use it for their work: formally or informally, directed or volunteer … they're pretty motivated to do it, and so I would say I think that the success rates are pretty good." Similarly, Officer 3 said:

> If there's an expectation that somebody's going to have their ML to be able to do their job … I should imagine they get on with it. But if there's no real drive … [they] kick it down the road and, "I'll get around to it, maybe, or maybe not. It's not a big deal." sort of thing.

Five participants said that some candidates had registered for the Mountain Leader qualification to develop their personal skills and that for some of these candidates passing an assessment was not important, Officer 4 said, "The ones doing it for their own skill improvement, it's not part of a definite plan … they're not so concerned if they complete or when they complete the award." However, Officer 3 did not believe that candidates attend a training course without any intention of going onto assessment but did think that some will decide not to continue:

> I don't transpire [sic] to this "doing the mountain leader training course for a skills course", to up-skill for an individual …. I can see how people

> would do it to start with, thinking they were going to progress to assessment, work out what were the demands upon them of attending an assessment, decide that we're going to call it a day there.

Provider 2 did not talk about candidates who only registered for the Mountain Leader qualification to develop their personal skills.

2.3.1.2.2 Regulatory Motives. Regulatory motives can be placed on a continuum from autonomous to controlled. Intrinsic-, integrated-, and identified regulatory motives are examples of autonomous regulatory motives, where behaviour is self-determined behaviour as the value of it is internalised. Whereas controlled regulation includes introjected- and external regulation, where behaviour is nonself-determined and the value of it may only be slightly internalised or not at all, thus controlled by external factors (Deci and Ryan, 2000). In these data, participants gave examples of candidates who had different regulatory motives and the influence that these had on candidates' likelihood of attending an assessment.

2.3.1.2.2.1 Autonomous Regulatory Motives. All seven participants said that those candidates who wanted to be outdoor instructors got to assessment. Officer 1 said, "[If] they're wanting to work in the outdoor sector they will naturally get [to assessment]." Provider 3 suggested that those with autonomous regulatory motives were more likely to get to assessment, "If you've got people that are thinking about a full-time career in the outdoors … they are going to be more inclined to follow the process through."

Another example of candidates having different types of participatory- and regulatory motives was seen in candidates who aspired to hold higher Mountain Training qualifications, of which the Mountain Leader qualification is a prerequisite for. Officer 2 said, "they'll tell you, 'I am doing this because I want to do my MIA.' " Participants suggested that these candidates were extrinsically motivated but had autonomous behavioural motives. Officer 3 gave supported this when describing his own experience of becoming a Mountain Leader, "I didn't even want to do my ML, I just wanted to go and be an MIA. I was only interested in that …. I was pretty flipping motivated to get through this thing as fast as I could."

Officer 4 suggested that candidates who aspired to hold higher Mountain Training qualifications wanted to complete the Mountain Leader qualification quickly to progress, "Folk that have got a definite plan for using their ML, like they want to become an IML [International Mountain Leader] or whatever either will pursue it in a shorter time frame."

2.3.1.2.2.2 Controlled Regulatory Motives. All seven participants talked about candidates who had controlled regulatory motives and suggested that these candidates were less likely to be assessed than those with more autonomous regulatory motives. When talking about which candidates attend assessments Officer 1 said, "If the school has sent them there because they're going to run a Duke of Edinburgh, then no. They won't do it." Provider 3 supported this and said that is because these candidates had not gained the necessary experience:

> We see a lot of people coming through with Duke of Edinburgh and Scouts who I'd say are pushed into it … they don't have the experience – the mountain experience as opposed to, sort of, hill and moorland experience – and it can be a shock. And then actually progressing through to assessment: they sort of realise, "Hang on." Yes, "I can't do this," or, "This isn't for me."

2.3.1.2.2.3 Intrinsic Regulatory Motives. Candidates with intrinsic regulatory motives also had intrinsic participatory motives, at least to attend training. Those who did not feel that they wanted to be assessed were intrinsically motivated to attend a training course but amotivated to complete the qualification. Provider 1 suggested that if candidates registered for the Mountain Leader qualification to develop their personal skills and found their training course inspiring then they were more likely to want to be assessed:

> If you run a good course, you enthuse them so much that there's no requirement on them to come back and do the assessment, but they actually want to do the assessment because they feel that it's a good challenge for their hobby.

This was supported by Provider 3:

> Quite a few who come on training courses and say, "Oh, I'm just doing this for a personal thing," actually really enjoy it, and then they go, "Oh, I'm going to carry on now and do the assessment, and actually this seems like a really cool thing."

Officer 4 suggested that candidates' motivation and self-efficacy can be influence by course staff:

> It's a combination, isn't it? Of helping them believe they can do it and helping them want to do it, to see value in completing, because a lot of folk come on training courses not being sure they need to do the assessment.

2.3.1.2.3 Negative Disconfirmatory Experience. All participants talked about disconfirmatory experiences that reduced candidates' motivation to attend assessments. However, three of the participants also provided evidence that not all candidates who have these experiences will drop out. Officer 2 proposed that all candidates will have at least one such experience, "I would be really surprised if they have never had a disconfirmatory experience." Officer 1 and Officer 3 go further and suggest that some candidates may become more motivated following a negative disconfirmatory experience. Officer 1 summarised the possible effects of negative events on getting to assessment by saying that, "[candidates] either do a U-turn and don't bother or they up their game."

Five participants gave examples where candidates were part way through the Mountain Leader qualification process and realised that it was not something that they either needed to or could do. Officer 1 gave the following example:

> Someone who … saw a Mountain Leader working, thought, "That's the thing for me," … and then once they started the process realised [that] actually there's a lot more to it than they were hoping and then become disinterested with how much experience they needed to gain from then on it, and then dropped off.

2.3.1.2.3.1 Negative Experiences at Training. Six participants suggested that in some instances a Mountain Leader Training Course itself could be a negative experience. When asked for an example of a disconfirmatory experience, Officer 3 said,

"[a disconfirmatory experience] might be just feeling they are well off the mark during a training course … that can be quite depressing … just not really nailing it on the training and then getting disillusioned."

When talking about candidates who felt less willing to attend an assessment Provider 2 said, "People say, 'It really put me off. The training course really put me off,' and that's a shame when you hear that because they say, 'It was just awful.' " Provider 2 repeated examples that candidates had previously given to them of reasons they had become less willing to attend an assessment:

> A lot of comments come, "Our training was worse than the assessment" …. "We never had any feedback. We were assessed basically" …. These people went on their training course and felt like they were beasted and battered and scrutinised like as if they were being assessed.

Six participants spoke about candidates who had not understood the purpose of the qualification when they had registered for the Mountain Leader qualification and once the candidates better understood the purpose of the qualification, they realised/decided that they could/would not complete it. Provider 3 explained that the training course had sometimes been the stimulus for candidates making that choice, "We definitely get [candidates] that are coming forward and then they do the training course and they realise it is just not for them, they are not going to be able to put the time and effort in."

Some of the candidates who decided that they could/would not complete the qualification following their training course may have done so based on incorrect information. Officer 1 said:

> We have had cases where someone has asked about experience [needed prior to assessment] and a provider has gone, "Well, in my view everyone needs to go to Scotland and go to the Highlands to gain experience" …. Suddenly people are going, "Oh, my God. I live in the South East …. If I have to go to Scotland that's a whole different ballgame."

Officer 1 went on to explain that the quality of information provided by training staff determined if it had a positive or negative influence on candidates, "The wrong

kind of responses [from training staff] can have an impact. Whereas the right answers might mean that people get the correct information and can then plan accordingly."

Officer 3 provided an example where candidates' perception of the course staff as role models might discourage them from completing, "I am sure there is nothing more disengaging than seeing somebody out of shape, out of currency doing a crap job on the hill. It is hard to engage with that."

2.3.1.2.4 Competing Influences. Five participants spoke about candidates who wanted to complete the Mountain Leader qualification but were not motivated enough to find the time to prepare for and then attend an assessment. There was some evidence that those who take longer to complete the Mountain Leader qualification will need more enduring motivation. Officer 3 said, "Sometimes I think momentum is everything." Officer 1 supported this:

> I think those who see it as the end goal take longer, and the more time that you put in between that training and assessment there are more variables of life that can get in the way that would then push that to the back burner.

When asked about candidates who were ready for assessment yet did not attend one, Officer 3 said that the Mountain Leader qualification is, "an easy can to kick on down the street if you're busy with other parts of your life." Officer 2 supported this idea of candidates having put their assessment off because they were busy with other things:

> Maybe they haven't turned up to assessment at that point because they haven't got the days, and said, "You know what, I haven't managed to get the days in, I'll leave it this year, I'll do it next year." That's fairly common …. There are just other things, life's got busy in other ways.

Officer 1 explained that following a training course some candidates realised that they would need longer than previously expected to complete the Mountain Leader qualification. For some of those candidates their motivation to complete the Mountain Leader qualification did not last:

> Where candidates lose focus is if they've found that the training course has brought lots of new skills to them that they haven't seen before, they

> start pushing back when their assessment time's going to be. I think once that goes beyond 12 months, they kind of come off the boil with their consolidation time because it feels like there's no urgency …. I think once they do that they're less committed, so making good use of their free time to consolidate and gain further experience becomes less of a priority, so the further that goal is the less a priority it becomes in their everyday life. Then that opens up lots of opportunity for life events to get in the way.

2.3.1.3 Barriers to Gaining Experience.

One prerequisite for a candidate to attend an assessment is having a minimum experience of 40 Quality Mountain Days (QMDs). Accruing 40 QMDs requires the investment of both time and money. All seven participants discussed reasons that candidates had not meet this prerequisite and thus did not attend assessments. All seven participants spoke about aspects of candidates' lives that prevented them from gaining sufficient experience to get to assessment. Officer 1 said, "If people can't get the experience they can't proceed." Provider 3 supported this by saying, "Location and time, I would say are the biggest two handicaps for people. So, if you don't live in the mountains and you've got a fulltime job and a family, really hard." When asked how different motives for doing the Mountain Leader qualification influenced a candidates' chances of completion Provider 1 said:

> Well, really, it boils back to, "Are they in a position to gain that experience to go forward to assessment?" That's the actual crucial thing, I think, more so than any one group where you go, "Yes, they're much more likely to do it."

Officer 3 supported this saying, "I think timing is critical, you have got to have the time to gain experience. You have got to have enough money in the bank to get through the process."

2.3.1.3.1 Lack of Time. Participants gave three main reasons that candidates felt they lacked time to prepare for their Mountain Leader assessment: profession, family, and doing other multiple qualifications at the same time. These other domains of

candidates' lives became barriers to completion for them as they were more important to those candidates than becoming Mountain Leaders.

2.3.1.3.1.1 Profession. All seven participants suggested that candidates whose profession allowed them time to prepare were more likely to be assessed than those whose profession did not. How a candidate's job is set up appears to be more important than what that job is.

An example of candidates in the same profession having different amounts of time to prepare is clearly illustrated amongst trainee instructors; five participants spoke about how different trainee instructor schemes influenced how much time candidates felt they had to prepare. When asked how being a trainee might affect a candidate's chances of completion Provider 3 said, "[Outdoor Centre A] and people like that with, some of their staff are very good at giving them time off, or sometimes even paid time to go and do a bit of personal development." And when talking about candidates from outdoor-activity centres Provider 2 said:

> If you're just given week after week of programmes that demand your time, working at low level, and the organisation is not giving time to develop their own skills … It's down to the company you're working for and it's down to the organisation. They're the ones who will decide what they need and how much time they've got available to release.

This was also evident in the five interviews where participants spoke about how being a teacher influenced a candidate's likelihood of attending an assessment. Provider 1 explained that teachers who felt that they only had their holidays to prepare for the Mountain Leader might have felt that they could not "fit it in" and that teachers available time is dependent on their school's view of the Mountain Leader qualification:

> I mean, schools can be helpful or not so helpful …. If the head teacher gets outdoor ed. and all the good things that spin out of it, then they can be very supportive. If the head teacher doesn't, then the teacher's kind of fighting them as well with all the other pressures: family, money and whatever.

2.3.1.3.1.2 Family. All participants said that candidates having family commitments would make them feel that they had less time to prepare, so were less likely to get to assessment. For some candidates, this was moderated by support from their family, allowing candidates to prepare for the Mountain Leader assessment instead of fulfilling their family commitments. When asked for examples of reasons people have given for not completing the Mountain Leader qualification Officer 2 said:

> Family. Family and work. Kids, or family circumstances, maybe elderly parents. That seems to be the main thing, or work commitments …. Sometimes they come back … they have resurfaced on the other side to say, "I am picking this back up again."

Officer 1 gave an example of a candidate whose family situation, and thus priorities, changed between training and assessment, which meant that they had not and were unlikely to complete the Mountain Leader qualification:

> Three years ago, I talked to her about doing the ML. She cracks on with doing that. She's done the training. She hasn't done the assessment. She's now had a kid, and it's almost totally irrelevant to talk to her about ML these days.

Officer 2 explained that candidates from different backgrounds will have different levels of family responsibility when talking about candidates from minority groups, "Sometimes when folk in other communities get involved in the outdoors there are religious, cultural and social pressures …. Family commitments come first, and it has a big impact on free time … suddenly your free time isn't free."

2.3.1.3.1.3 Multiple Qualifications. Some candidates also work towards other qualifications at the same time as the Mountain Leader qualification. Five participants suggested that working towards multiple qualifications at the same time has a negative impact on the time available to candidates and thus their likelihood of attending an assessment. Officer 2 explained that working towards multiple qualifications at the same time made it harder to do one well:

> [Candidates] who tried to then spread with paddle sports and that really suffered …. You have to have a bit of a focus …. You have to decide which

> one it is you are going to do. Unless you are one of these really rare people who's brilliant at everything.

Provider 3 suggested that working towards multiple qualifications at the same time may be detrimental to a candidate's chances of attending an assessment because of changes in their regulatory motives:

> Sometimes they're trying to do quite a lot of tickets all at the same time and it can become a chore for them, and it's almost like a hoop that they feel they need to jump through as opposed to actually enjoying the process … I think a lot of them find it really hard to put the time in.

Officer 4 also recognised that working towards multiple qualifications at the same time may limit the amount of time that candidates can gain experience in but suggested that there might be some advantages to this as well:

> [Trainees] might be preparing for other things at the time. But equally, they're in a particular phase of their life and mind-set, which is award focused. So, therefore, they will be quite good at preparing for assessments and more likely to have access to other people that have got MLs that can help them.

2.3.1.3.2 Location. Six participants discussed how the place where a candidate lives influences how easily they can accrue QMDs. It is harder for candidates who live further from the mountains to accrue QMDs as they must both travel for longer and often feel that they need to take block of time off to get to the mountains. Officer 2 explained that candidates living in Scotland could gain QMDs "in a day rather than two days" because they did not "have a day's travelling to get there and back." This was supported by Provider 1 who said, "People for whom the mountains are a long way away: by definition, it's going to be harder because they've got to have the time and the money to get themselves there." Provider 1 went on to say, "They're going to do it more as bunches of days, so they're quite likely to do multi-day expeditions …. Whereas, the people who live closer can do it weekend and weekend, once a month on a Sunday."

In addition, candidates living further from the mountains will face a higher financial cost. For some candidates, this can seem beyond their means, Officer 1 said:

> The financial cost of gaining the experience is a massive challenge. When you're talking to someone from the South East, telling them they need to get up into Snowdonia and The Lakes, or The Highlands, on 40 occasions, they start going, "Bloody hell. I can't afford that."

2.3.1.4 Social Support.

2.3.1.4.1 Development Plans. All seven participants said that it was important for candidates to leave their assessment with an understanding of what they needed to do to prepare for an assessment (i.e., have a development plan). When asked what the most important part of support was for candidates, Provider 2 said:

> Once they've got onto the training a really good training course, which makes it clear to the candidates what it's all about, and then directs them the right way. You need to individually debrief people and get to know what their personal needs are … A generic debrief really sometimes doesn't cover it thoroughly enough for individuals.

When asked what influence they thought the post-training debrief has on candidates Provider 3 suggested that it could have a profound impact on candidates' expectations:

> It's a really important chat … it's really common on a debrief when you sit down with somebody and say, "That was an awesome performance. All you need to do is pad this logbook a bit, and you could come forward for assessment really quickly." They sit there and go, "But I was thinking about doing it in four years' time." and you're like, "What? You could do it next spring, no problems at all" …. You can have a big impact.

However, Officer 3 explained that providing individualised feedback can be at odds with preventing training courses feeling like an assessment, an issue highlighted above (see [Negative experiences at training]), "I don't believe that candidates should feel they're under any sort of assessment process while on the training course. Once you have a formalised one-to-one debrief it can feel like an assessment." Provider 1 suggests that it is possible to provide individualised feedback without making candidates feel that they have been assessed:

> My debrief is actually getting them to tell me what they think they need to do rather than me telling them what they need to do, because I would've had to assess them somehow to do that …. I'm asking them to self-assess and tell me what they think they need to do to get to the assessment.

2.3.1.4.2 Time Support. As shown above (see [Profession] and [Family]) some candidates felt that they did not have enough time to prepare for a Mountain Leader assessment. However, different candidates with the same demands on their time can feel differently about the amount of available time they have. One reason for this is that some candidates are supported by their employers and families. When asked what sort of support candidates might look for Officer 2 said,

> Having the support of their family is going to be absolutely paramount …. Having support from family to free up time and then actually having the time both from family and work that coincides with the others …. It is an acknowledgement within the family that [the Mountain Leader qualification] is important to the person. The ones who have succeeded against the odds have had that support. That's been really obvious.

Employers are another source of time support for some candidates. When talking about support candidates received with practical matters, Provider 1 said, "Some of them are in organisations and centres where the management are on the ball enough to allow them development time."

2.3.1.4.3 Financial Support. Six participants spoke about candidates who had received financial support. In some instances, this was essential to candidates' progression to assessment. Provider 2 said that, "A lot of people wouldn't be able to do ML if they didn't get financial assistance" and went on to say, "However, participants also suggested that financial support will only benefit candidates if they are also sufficiently motivated to complete the ML." Officer 3 said:

> In my experience, those [whose] pathway has been paid for or financially supported, they don't really seem to engage with the actual role of taking

> responsibility for a group in the mountains …. Heavily subsided or full payment I tend to find they don't get a good solid engagement and on occasions people just don't turn up because there's no engagement at all.

When talking about candidates who want to use the Mountain Leader qualification for work Provider 1 supported this paradox associated with financial support, saying, "they're pretty motivated to do it, and so I would say I think that the success rates are pretty good for that, particularly if they've paid for it."

2.3.1.5 Reengaging Later in Life.

Five participants discussed candidates who had disengaged with the Mountain Leader qualification and then but reengaged with it later in life. Provider 2 gave an example where candidates had an enduring motivation to become Mountain Leaders but had not completed the qualification because they were busy with other aspects of their lives:

> They start the process when they were young, free and single. They meet somebody, get married, have kids, they don't do it for years and years and years. Then they come back to it. It's something they've always wanted to do.

Provider 1 also suggested that changes in family circumstances can be the reason that candidates reengaged with the Mountain Leader qualification:

> The Scouts, the Guides and the D of E are often the kick-start to get people back into it again because they've suddenly found that their kids are actually at that stage …. Then, they want some formal training on top of that.

Provider 3 suggested that retirement might also provide candidates with an opportunity to reengage, "[Candidates] who did their training a long, long time ago and then their career is coming to an end …. They'll reengage as well."

2.3.1.6 Redirection to Lower Qualifications.

Five participants suggested that after Mountain Leader training some candidates decided to pursue a lower qualification instead, Officer 1 said, "They can't put [the

Mountain Leader qualification] as the priority in their life, so they may drop back to the Hill and Moorland Leader or the Lowland Leader course as a more achievable objective." This was supported by Officer 2 above (see [Family]) and when talking about candidates who have struggled with the Mountain Leader training course, "We get a reasonable number that then convert to Hill and Moorland Leader …. They decide that they are going to do that, because that is a shorter assessment and less intensive."

It is unclear how this will ultimately influence getting to a Mountain Leader assessment. For some candidates this lower qualification will suit their needs and they will not continue with the Mountain Leader qualification but for others, completing the lower qualification becomes another step in the process of becoming a Mountain Leader. When talking about training debriefs, Provider 2 said, "Sometimes, we would advise somebody to go and do the Hill and Moorland assessment …. They worked really hard to get the Hill and Moorland … then eventually, after a couple of years, they've done the ML assessment."

2.3.2 Passing

2.3.2.1 Ability.

2.3.2.2 Resilience.

All seven participants spoke about the need for candidates to be able to recover from setbacks in order to pass an assessment. Indeed, mountaineering and leading others in the mountains can often be a stressful experience; when talking about the perception that Mountain Leader assessments are stressful Officer 2 said:

> I've been in really shit situations with clients, and there are no assessments as stressful as when things aren't going well when you've got real people there. So therefore, if you haven't got the mechanisms and strategies to deal with assessment nerves and stress, then actually you're not really going to cope well when things go horribly wrong with a group.

Five participants felt that at some point during a Mountain Leader assessment it was inevitable that candidates would make a mistake, even the most competent, and how candidates deal with those mistakes is important, Provider 2 said:

> I always say to people, "It's very, very unlikely that you will not make a mistake because making mistakes is part of it. We all mistakes so if you make mistakes, learn from them and move on. Be positive with it." It's very difficult to be assessed or watched for five days without making an error of some kind, and I said, "When you make mistakes you will blow those mistakes right out of proportion because you'll be hard on yourself. You'll have your moment where you've got to get through that. It'll happen to everybody probably. It's really unlikely to get through this week without making mistakes so you mustn't let these mistakes get you down. You've got to keep coming back. When you rectify mistakes, it tells us [assessors] a lot about you. It's no good you getting disoriented a bit and throwing the towel in because what we want to see you do is get disoriented and sort it out because we've got to think if you were with a group in that situation, what would you do? Throw the towel in or would you sit down, concentrate, re-orientate, think about it and sort it out?"

Five participants spoke specifically about experience building resilience as candidates with more experience are more likely to have dealt with setbacks as part of that experience. Provider 1 said that, "we want people to be able to bounce back, that's perhaps a product of experience, and greater experience means that they're more likely to have to do that because if you're in the mountains long enough things go wrong." Officer 2 explained how Mountain Training have incorporated this phenomenon into the concept of a QMD:

> One of the reasons for creating the concept of Quality Mountain Days, I try and explain to people, it's not just any old day, it's challenging days. And the idea really … is to develop resilience. So, you go into situations where you're challenged on a wide variety of levels, both technically, and physically, and mentally. And if you have loads of quality [mountain] days, where all of these elements are taxed and challenged. Then you come through it, or sometimes you don't but you learn from it, you'll develop resilience. And you're used to dealing with adversity. And that's what a Mountain Leader, I guess, at the end of the day, when push comes to shove, that's what they've got to deal with. But I think folk

> short-circuit the quality mountain day experience. So, any day is a quality mountain day, and therefore that resilience isn't necessarily as great.

Officer 3 explained that candidates' experience outside of the mountains may also help them to cope with adversity:

> Candidates' experience and abilities that is important… I guess we touched on yesterday, it's that ability, the resilience and robustness of the candidate. That might be born partly from their mountaineering experience, it could also be lessons they've learned in other aspects of their lives that they can very easily transfer to coping with adversity in that mountain context.

2.3.2.3 Staff Behaviour.

2.3.2.4 Experience.

All seven participants discussed the importance of candidates having experience in order for them to pass a Mountain Leader assessment. Above, experience has been related to increased levels of confidence (Section 2.3.1.1) and resilience (Section 2.3.2.2), in addition influences on candidates' ability to gain experience have been discussed with reference to candidates getting to an assessment (Section 2.3.1.3). Participants discussed three facets of experience that were important when considering the outcome of an assessment: quantity, quality, and variety. The relationship between each of these facets and candidates' performance at assessment are discussed below.

2.3.2.4.1 Quantity. All seven participants spoke about how the quantity of experience a candidate has influences the outcome of their assessment and suggested that, in general, candidates with more experience would be more competent and therefore more likely to pass. When asked about the performance of candidates at assessment, Provider 1 said that, "A lot of it is about experience, that you build up … by being in the mountains and having done all that stuff." Provider 2 supported this, saying, "People's performance is more down to the level of experience and the amount of preparation they've done for that week."

Having 40 QMDs is a prerequisite for passing a Mountain Leader assessment and all seven participants emphasised the fact that 40 QMDs is the minimum, explaining

that having the minimum experience is not always enough for candidates to demonstrate competence. Officer 1 said:

> If we have a candidate that reads the assessment criteria and does the minimum, is aiming for the minimum, so that is visiting 3 areas, getting 40 quality mountain days, then typically you find that they're struggling to make the right decisions and adapt to different scenarios and different places.

However, four participants suggested that for a minority of candidates, 40 QMDs were more than enough. Officer 3 said:

> It's a bit of an issue for course directors sometimes, where during the practical assessment, the candidates show evidence satisfying all the competencies, as such. Therefore, they want to pass the person. But when they look back in their DLOG, they find that, actually, they've got less than 40 Quality Mountain Days, so it's a logbook deferral. Which seems a bit weird to me, that, if they've shown evidence that they can do the job on the hill. Almost, the logbook becomes less relevant. But it's what we do. I guess they're few and far between, those. Regularly, poor performance goes hand in hand with a weak logbook.

Three participants spoke about the relationship between the quantity of experience that a candidate has and their level of confidence. All three suggested that experience develops confidence and confident candidates usually perform well. Provider 2 explained that:

> Loads of people turn up really confident because they're really good and they've got really strong logbooks and they are confident in their skills. That is from the word go. They have done so much preparation, so much, they are so used to what we're going to do now this week, so they just cruise it …. At the end of the day, we're going mountain walking and if they do a lot of mountain walking and they're comfortable with that then they're just good aren't they and they're confident.

Four participants described reasons that candidates had not been able to gain as much experience as they would have liked to before being assessed. Section 2.3.1.3 discussed barriers to candidates gaining enough experience to get to an assessment; for some candidates, these barriers did not stop them getting to an assessment, but they did prevent them gaining the experience that the candidates would like to have prior to being assessed. Officer 1 gave an example of this and the effect it had on the candidate:

> They made the decision to book [an assessment]. Then a life event got in the way, didn't allow them to consolidate as much as they wanted to, but they still went through with it, to try and give it a go. Then I think when they arrive, having not done the preparation that they knew they needed to do, day one you do introductions around the room, and then they've got other people in the room who appear very well prepared, that then knocks them back.

However, quantity of experience on its own is not enough. Candidates also need to have sufficient quality and variety of experience, Provider 1, "If they've got a huge mountain experience, yes, it will serve them better. It can't help but not, really, and the wider that experience almost the better."

2.3.2.4.2 Quality. All seven participants spoke about the importance of candidates having experience that was of suitable quality, as it is possible for candidates to gain 40 QMDs, but not develop their skills as the experiences would not have been challenging enough.

2.3.2.4.2.1 Weather All seven participants spoke about the importance of candidates having experience in bad weather. Officer 1 said that, "There's less value in lots of good weather days in terms of gaining experience." Provider 1 explained that a lack of experience in bad weather prior to assessment could leave candidates unprepared, "In the past, [candidates] have come adrift on assessments because they've not been out in bad weather." Officer 2 supported this, and explained how they thought having experience of bad weather developed candidates' resilience:

> We could tell the ones who'd been out [in] really crap weather, really stormy, horrific weather; natural propensity was not to go out in it. Well

> actually, some did. And when they all came to do the assessment, you knew the ones who'd been out in it, because they could deal with it. And the ones who couldn't just went to pieces. And that was resilience. And you knew that they'd paid their dues, and it had been probably really tough. And hats off to them, they'd put themselves into quite unpleasant, and probably quite dangerous situations, potentially, and they've come through it. And then they've paid their dues when they came on the assessment, and they were facing similar conditions. "Well, this is… I've done this before." Whereas the ones who avoided that, because it was unpleasant, didn't. So I think, in a way, we rely on that resilience to be developed, just by doing experience.

2.3.2.4.2.2 Off the Beaten Track All seven participants spoke about the importance of candidates gaining experience "off the beaten track." To pass a Mountain Leader assessment, candidates must be able to navigate in a variety of mountainous terrain. However, it is easy for one to spend time in the mountains, but never venture from an established path; Provider 1 said, "Let's face it. Most mountains you walk up you walk up the path …. When do you ever go up a mountain that you don't walk up the path? You would have to deliberately not walk up the path." Provider 2 explained that candidates may have lots of experience, but not in appropriate terrain, which results in their skills not being at the standard required at assessment:

> They might have done 100 mountains but every mountain they've done is on a major footpath, for example, and they've never really ever gone into any steep, complex territory so they're struggling as soon as they're in that territory and you've got to test people in that territory just in case they go there or they end up being there or they choose to be there for any particular reason and that is the standard isn't it? Sometimes people struggle off the beaten track, they can't get their head around that side of things and you look at their logbook and they've done a lot of footpaths. But then again, why shouldn't they? If you go up Scafell, you're going to go up a footpath aren't you? If you go up Snowdon and you've never been here before you're going to go up the footpath. It's logical so it's pretty

normal I think.

2.3.2.4.3 Variety. All seven participants discussed various ways in which candidates' experiences varied and the effect these differences had on their performance at assessment. In general, more varied *relevant* experience was related to better performance, however, experience that was not relevant was not. Participants spoke about four different aspects of variety of experience: variety of QMDs, experience of other assessments, climbing and mountaineering experience, and international trekking experience.

2.3.2.4.3.1 Variety of QMDs Five participants discussed the variety of QMDs that a candidate may have, and all five participants suggested that the more variety in QMDs that a candidate had the better. Indeed, Officer 4 suggested that the variety of QMDs a candidate has is more important that the quantity:

> I guess the type of experience they've been getting is more important than the absolute volume. A lot of folk do quite repetitive stuff, feeling they're gaining good experience, where actually, if it was condensed into fewer but more varied; they would be learning a lot more.

More specifically, Provider 1 suggested that a benefit of gaining experience in different geographic locations was that it has exposed candidates to a wider variety of terrain:

> If somebody had 40 days only in Snowdonia, compared to 40 days where that was spread across Snowdonia, the Lakes, the odd days in the Peaks, and the West and East in Scotland, would the second one be better? Yes, of course they would, because they've just got a greater experience of different types of terrain.

Officer 3 supported this and also suggested that it was important that QMDs were gained in areas unfamiliar to the candidates, "I think if somebody just goes out and does loads of varied days in the mountains in all sorts of weathers in places they don't know, then I think they will get through."

2.3.2.4.3.2 Experience of Other Assessments Five participants spoke about the benefits of candidates having previously attended assessments, especially if the assessments were similar in nature to a Mountain Leader assessment (e.g., practical, continuous in nature). When asked for examples of the types of candidate who are confident that they will pass when arriving for their assessment, Provider 3 said:

> Anybody that's been through a similar process already, so maybe they've done their paddling qualifications, so a similar outdoor qualification, will have a better idea of what to expect. Military personnel who have already been through a military process have a better expectation of that. And, maybe people who are further along in professional careers that require some sort of continual assessment, you know, so they're just used to being looked at and being assessed and having to revalidate with qualifications and things like that, they tend to be a little bit happier in that environment as well.

In contrast, it was suggested that candidates who were less familiar with assessments were more anxious. Indeed, Provider 2 said that, "when I did my first training and assessment, I would say that I was probably a bag of nerves and I wasn't sleeping properly the night before it started." Officer 2 suggested that candidates who were less familiar with assessments sometimes behaved in ways that were unusual for them:

> You also get people that might be a little bit older who haven't been assessed for a very long time …. When you review it with them they go, "I don't know why I did that. I wouldn't normally do that. I thought that might be what you wanted to see" …. Whereas an outdoor instructor who's going through multiple qualifications is getting very used to peer review, receiving feedback, being trained, being assessed, and they're enjoying the process. That is going to make a difference to how people then do things throughout their assessment week.

2.3.2.4.3.3 Climbing/Mountaineering Experience Five participants discussed the influence of climbing and mountaineering experience on candidates'

performance at assessment, however, their views were somewhat nuanced. All five participants felt that climbing and mountaineering experience could benefit a candidate's performance, as they were more likely to be confident and proficient in their mountaineering skills. Officer 2 said:

> Candidates with a much more broader mountaineering experience, they're generally more relaxed because they've been to lots of different environments, they've made a lot more decisions, they've had different circumstances and that makes it more adaptable. So their approaches to the technical skills it's more common sense approach I guess rather than a clinical, "This is how I need to navigate to get out of here." Yes, broader experiences is something that I think would really help someone towards getting through an assessment successfully.

Whilst the potential benefits were recognised, four participants (including all three providers) described candidates with climbing experience being overconfident or not appreciating the difference in climbing and mountain walking. Provider 3 said:

> There can be a negative side of things: People that are on climbing quals – you know, SPAs and things like that – can be overconfident in what they think is suitable terrain to jig people around on; I see that quite a lot. And yes, that's not nice when you see that. And I warn them. (Laughter) I do give them a heads-up that this is their Summer ML and not their SPA and they need to be behaving appropriately. So unfortunately yes, I've seen some quite poor performances because of overconfidence because they think they're a climbing instructor and there's a lack of appreciation of the difference between the two.

Interestingly, Officer 4 linked this inability to differentiate the context to experience, with those who were less experience being less able to make appropriate judgements, "If your experience is quite limited, it's quite hard to know, to sort of sift through that difference to ML rope work, climbing rope work and the two can get a bit confused."

2.3.2.4.3.4 International Trekking Four participants discussed the influence of international trekking experience on candidates' performance at assessment. Whilst this experience would prove somewhat useful, it was something that could compliment QMDs, not replace them, as some skills that are required for the Mountain Leader are unique to the UK, an example of this is navigating away from paths, the importance of which is discussed in Section 2.3.2.4.2.2. Officer 3 said:

> The reality is, unless they've actually hung out in the British mountains a bit, they don't perform very well. Irrespective of how compatible they think the [non-UK] environment is to the UK, it's very unique. So experience in the UK mountains is the most important thing, and developing a personal skillset.

2.3.3 Reassessment

Candidates who do not pass their initial assessment may or may not return to be reassessed. Whilst the interview was not designed to answer questions about what factors influence if candidates return to be reassessed, the semi-structured nature of them meant that some data emerged that provides some insight into this. However, these results are not as clear as those in previous sections.

2.3.3.1 Understanding the Original Result.

All seven participants spoke about candidates either understanding and accepting their original assessment result or not. Participants suggested that candidates who understand and accept their result are in a better position to decide if they want to continue with the Mountain Leader qualification and if they do, understand what they need to do to pass a reassessment.

2.3.3.1.1 Preparing for Reassessment. Four participants spoke about candidates who realised that they were below the standard and then went away to prepare for reassessment, Provider 3 said:

> You'll get lots of candidates who get deferred on their navigation, and are like, "Urgh," then they go away, sort themselves out, come back for

> reassessment, and at the reassessment process they go, "I definitely wasn't good enough, and I've gone away and done all this stuff, I know realise I'm a much better navigator than I was before."

2.3.3.1.2 Disagree With/Do Not Understand Result. Three participants spoke about candidates who did not either agree with or understand their original assessment result. Officer 4 gives an example of why candidates might not agree with their result:

> I guess there's a danger that [candidates] don't fully understand, the reasons for having been deferred. If they're pinning it on isolated, you know, isolated mistakes that they've made or errors. Maybe they haven't grasped that it's a pattern that's emerged.

2.3.3.1.3 Consequences of Not Understanding/Accepting the Result. Officer 1 explained that candidates who felt that their result was unfair would do one of two things:

> They would literally finish that assessment. Get the result they didn't want to hear. Then they will either do one of two things, complain, or just get annoyed, and try and book onto the next earliest assessment they can. They don't believe they need to retrain. They believe they need to just be assessed again. Then they go to that next assessment and, hey presto, the same result …. Because nothing has changed. Unless it is about the assessor/candidate relationship …. If it's about the system, rather than about the assessor, then if they go on to the next assessment they will just get the same result again.

2.3.3.1.4 Reasons for Not Understanding/Agreeing. Provider 2 suggested that clashes between candidates and staff are not uncommon:

> You always have people complain about something or somebody, sometimes, about situations they're in. You get a lot of info when people are being reassessed because they'd been deferred so you run a re-assessment and because they're being deferred, they start telling you

> why they think they shouldn't have been deferred and then they start slagging off providers and organisation.

Provider 2 goes on to explain that in some instances these clashes can be highly charged, "I have heard stories of people saying, 'I nearly punched him. I nearly hit him. In fact, we all did. We all felt like turning round and hitting him.' That's not good, is it?"

2.3.3.2 Nerves.

Three participants said that candidates who present for reassessment are nervous, Provider 2 said, "Everybody who turns up for a re-assessment is full of nerves. They're very nervous when they start". Two of these three participants suggest that for some this nervousness can be so extreme that it manifests itself with physical symptoms. Provider 1 gave an example:

> Some people are literally sick with worry on ML assessments. I mean, I remember doing a reassessment for this one guy and he confessed afterwards that, just before we'd met up, he was throwing up because he was that nervous about doing this.

Provider 1 supports this, suggesting that it is not a one-off occurrence:

> They are really nervous and when you meet them you've got to really make sure you calm them down and you've got to try to create a really relaxed atmosphere before you set off because they're shaking some of them. They're nearly sick.

This nervousness may be in part due to their experience on their original assessment. Provider 2 provided the following insight in candidates' original experience of assessment, "[Candidates] felt like [assessors] were quite harsh and quite lacking in any form of feedback or lacking in any form of empathy, which made them feel very uncomfortable, which made their performance even worse"

2.3.3.3 Redirected Towards a Lower Qualification.

Some candidate will not return for a reassessment because following their initial assessment because the assessment staff have redirected them towards a lower level

qualification, as the assessment staff feel that would be more appropriate for them. Whilst only Provider 2 spoke about candidates being directed towards a lower level qualification rather than reassessment this redirection is also spoken about in Section 2.3.1 Getting to Assessment. Provider 2 said, “When you get people like that, we advise them to do Hill and Moorland …. You will get that candidate who will be better off, definitely, doing Hill and Moorland.”

2.4 Summary and Concluding Remarks

Chapter 3

Key Discriminatory Factors

3.1 Introduction

From the results of Study 1 it was clear that there was no single factor that determined whether or not a candidate would complete the Mountain Leader qualification. Instead the results suggested that both the main effects of, and interactions between, a myriad of factors were important. Study 1 collected data from those who were involved with the organisation and delivery of Mountain Leader training and assessment courses. To our knowledge, Study 1 represents the most in-depth investigation of factors influencing the completion of the Mountain Leader. However, it does not include any data from candidates themselves, nor does it test whether these factors do actually influence completion. Therefore, to develop a more complete view of the factors influencing the completion of the Mountain Leader qualification, this chapter reports the findings of three studies that collected data for these factors from candidates who had registered and attended a training course for the Mountain Leader qualification.

3.1.1 Pathways to Expertise

The development of expertise and the pathways to it are of interest in a variety of domains. Historically, most studies have examined the impacts of specific factors on the completion of a training pathway (e.g., delays in completing PhDs; van de Schoot et al., 2013b). However, these single-variable approaches fail to acknowledge that there may be vast differences between individuals in their pathways to expertise. As such, a growing number of researchers now recommend that multi-disciplinary approaches should be

adopted for identifying the potentially complex interactions that influence talent/expertise development (e.g., Abernethy, 2013; Güllich et al., 2019; Johnston et al., 2018; Pearson et al., 2006; Rees et al., 2016; Schorer and Elferink-Gemser, 2013).

Developing expertise is likely to be the result of complex interactions between a variety of developmental factors (e.g., practice and training, personality traits, motivation, social support; Baker and Cobley, 2013; Gagné, 2004; Johnston et al., 2018). Further, different factors will be more salient at different points of development pathways (Rees et al., 2016). Based on these principles, a number of projects in the elite sport domain have recently begun to explore the most important factors in the development of athletes in elite pathways (e.g., Güllich et al., 2019; Hardy et al., 2017; Jones et al., 2019; Jones, 2019; Jones et al., 2020; Rees et al., 2016). These projects have used a mixed-methods approach, collecting both rich qualitative data from athletes and making use of state-of-the-art machine learning techniques to identify sets of variables, whose main-effects and interactive-effects were able to discriminate between athletes at different performance levels. The methodology for this chapter is based on these studies, however, we used the results of Study 1 to help us develop a quantitative survey tool, so that we could collect, and subsequently analyse, data from more candidates than we would be able to if we adopted the interview based approach of those studies.

3.1.2 Chapter Structure

From the results of Study 1 and a workshop with Mountain Training,[1] we identified 168 factors that were deemed as potentially important to the completion of the Mountain Leader qualification, which were operationalised as 529 individual variables. The findings in Study 1 also made it apparent that completion may be better considered as two separate components with different factors being most relevant to each: getting to an assessment having attended a training course and passing that assessment. Mountain Training do not set a maximum duration between a Mountain Leader training and assessment courses.

[1]This workshop involved a presentation of the results of Study 1 to the Mountain Training council, followed by a series of break-out focus groups. The council is made up of 25 members, each of whom represent different stakeholders. The break-out focus groups asked questions to help the research team understand firstly, if the results of Study 1 resonated with the council and secondly, if they felt that there was anything that may be relevant, but had not already been identified through the literature review and interviews.

Figure 1.3 shows that the median duration between training and assessment for all candidates trained between 2009 and 2018 was 1.13 years and the mean duration was 1.58. Based on this and the time constraints of this project, in this study we operationalised getting to an assessment having attended a training course as getting to an assessment within 18 months of training. Furthermore, female candidates are less likely to be assessed than male candidates ($\chi^2(1, n = 15433) = 47.33, p > .001$), therefore we decided to examine separately the factors influencing female and male candidates getting to assessment within 18 months of their training course

There are four potential results for candidates who get to an assessment: pass, deferral (candidates will need to be reassessed on part(s) of the syllabus before qualifying), fail (candidates will need to complete a full reassessment), and withdrawn (candidates who do not complete the five day assessment course will need to attend a full reassessment). In this study, we were primarily concerned with candidates passing or not passing their first assessment, therefore we grouped the three non-pass results into one, rather than considering the non-pass results separately. The analyses in Chapter 1 did not indicate differences in the pass rates for female and male candidates, therefore we included both sexes as a single group with sex as an additional factor in the analysis for passing first time.

The aim of the present chapter was to identify variables influencing completion of the Mountain Leader qualification. In Study 1, we identified potentially important variables, from which we created a survey tool of reasonable length (we estimated that it would take approximately 20 minutes to complete) to collect data for these variables (see Appendix B). In the present chapter, we used the survey tool to collect data from candidates in order to identify important discriminatory variables for each of the following classification problems:

1. Male candidates who are assessed within 18 months of their training course from those who are not.
2. Female candidates who are assessed within 18 months of their training course from those who are not.
3. Candidates who pass their first assessment from those who do not.

The remainder of this chapter is structured as follows. First, there is a brief overview of relevant constructs, then there is a general methods section that describes

the data collection and analytical method for the three studies. This is followed by three studies, one for each of the classification problems listed above. Each of these studies draws participants from the aforementioned data collection based on their sex and progress through the pathway—any deviation from the general method is described. Finally, there is a general discussion of the three studies where overarching themes are discussed.

3.1.3 Relevant Constructs

Given the number of constructs it is beyond the scope of this chapter to provide a detailed literature review of each construct. However, in order to aid the readers understanding of the relevance of the constructs we have grouped the relevant constructs into several domains and provide an overview of the domains and the rationale for their inclusion in this project in Table 3.1.

Table 3.1: Overview of the theoretical domains included in the survey tool and the rationale for their relevance.

Construct	Rationale
Big Five	The Big Five model of personality (McCrae and Costa, 1987) is widely used when considering individual differences (Allen et al., 2013). For example, individuals who are more conscientious will persevere and be more hardworking and ambitious, therefore may be more likely to complete the Mountain Leader qualification than those who are less conscientious. As another relevant example, extraversion has been associated with effective leadership (Judge et al., 2002) and decision making (Hardy et al., 1996).
Resilience	Higher levels of resilience are associated with positive outcomes, including overcoming adversity (Smith et al., 2008). Further, the results of Chapter 2 suggested that it was important that candidates could deal with setbacks to become a Mountain Leader.
Intention of being assessed	The theory of planned behaviour suggests that intentions are the best predicator of behaviour (cf. Ajzen, 1991; Ajzen and Madden, 1986). Several studies have found evidence that intentions have been shown to predict behaviour (e.g., Armitage and Conner, 2001; Hagger et al., 2002).
Expected time to assessment	The results of Chapter 2 suggested that candidates who intended to be assessed sooner after their training course were more likely to be assessed.
Personal projects	There is evidence that goal importance influences goal progress and that it also moderates the relationship between self-efficacy and goal progress (cf. Beattie et al., 2015). Further, it was suggested that those who had multiple goals were less likely to be assessed as they would be committing resources to attaining other goals.

Table 3.1: Overview of the theoretical domains included in the survey tool and the rationale for their relevance. *(continued)*

Construct	Rationale
Understanding of the qualification	The results of Chapter 2 suggested that it was important candidates understood the purpose of and the standard of the qualification. It was suggested that candidates who were less certain of the purpose of the qualification may attend a training course to find out more about the qualification and then discover it was not what they needed to do, and those who were less certain of the standard would find it more difficult to be confident.
Socio-demographics	The results of Chapter 2 suggested that there were socio-demographic variables were important to consider when understanding why candidates do or do not complete the Mountain Leader qualification.
Available time	The importance of candidates having enough time available to become a Mountain Leader was highlighted in Chapter 2, as those who did not have available time would be unable to prepare for an assessment.
Access to the mountains	The results of Chapter 2 suggested that candidates who have better access to the mountains are more likely to be able to gain experience and that living further from the mountains may be a barrier to completion.
Participatory motives	The results of Chapter 2 suggested that candidates with extrinsic participatory motives were more likely to be assessed than those with intrinsic participatory motives.
Regulatory motives	Self-determination theory suggests that autonomous forms of motivation are better for prolonged engagement and more robust in the face of adversity (Deci and Ryan, 1985a; Ryan and Deci, 2017, 2019). This was supported by the results of Chapter 2.

Table 3.1: Overview of the theoretical domains included in the survey tool and the rationale for their relevance. *(continued)*

Construct	Rationale
Self-efficacy	Self-efficacy theory suggests that, if sufficiently motivated, self-efficacy will be the primary determinant of their performance, how much effort they will put in, and how long they will persist—particularly in the face of adversity (Bandura, 1977, 1982, 1997). Higher levels of self-efficacy have been associated with higher levels of goal progress (Sheldon and Kasser, 1998), task engagement (Caraway et al., 2003; Walker et al., 2006), goal commitment for self-set goals (Locke et al., 1984; Locke and Latham, 1990), and on-task effort (Bandura and Cervone, 1983). The results of Chapter 2 suggested that candidates would need to feel confident in their skills before they would go to an assessment.
Self-guides	The results of Chapter 2 suggested that candidates would have a "threshold" of confidence that they would need to surpass before they were assessed. It was also suggested that this threshold would vary with age, gender, and personality.
Self-efficacy discrepancy	Self-discrepancy theory (cf. Higgins, 1987) would suggest that greater discrepancies between the actual self and self-guides would lead to greater motivation as one would try to reduce the discrepancy.
Training staff coaching behaviour	Coaching literature describes a number of benefits on a variety of outcomes, for example, performance/skills, well-being, coping, work attitudes, and goal-directed self-regulation (e.g., Theeboom et al., 2014; Weinberg and Gould, 2014). The results of Chapter 2 also suggested that the behaviours of the training course staff were influenced candidates in several ways, including having the potential to foster more autonomous forms of extrinsic motivation.

Table 3.1: Overview of the theoretical domains included in the survey tool and the rationale for their relevance. *(continued)*

Construct	Rationale
Life changes	There is evidence that goal importance influences goal progress and that it also moderates the relationship between self-efficacy and goal progress (cf. Beattie et al., 2015). Further, it was suggested that those who had multiple goals were less likely to
Negative experiences	Negative events were identified as important in Chapter 2, it was suggested that experiencing these events would reduce a candidate's motivation be assessed. In addition, self-efficacy theory would suggest that disconfirmatory events would reduce an individual's level of self-efficacy (Bandura, 1977, 1982).
Additional training	The results of Chapter 2 suggested that candidates who received additional training after their training course would be more likely to pass their first assessment.
Preparation for assessment	Preparing for an assessment will provide candidates with opportunities to have mastery experiences, which could increase their levels of self-efficacy (Bandura, 1977, 1982). In addition, candidates are required to have a minimum amount of experience before attending their assessment course; most candidates will need to gain additional experience after attending a training course. The importance of the quantity, quality, and variety of this preparation are highlighted in the results of Chapter 2.

Table 3.1: Overview of the theoretical domains included in the survey tool and the rationale for their relevance. *(continued)*

Construct	Rationale
Social support	Social support has been shown to have a number of benefits, including helping individuals deal with pressure (Freeman et al., 2011, 2014), which would likely be beneficial to candidates preparing for an assessment. In addition to this, the provision of tangible support was discussed in Chapter 2 as it could help them find the time to prepare for an assessment.

3.2 General Method

3.2.1 Participants

We contacted all candidates who had attended their first Mountain Leader training course in 2017 or 2018, inviting them to participate in the study (N = 2,867). One thousand and thirty candidates started the survey and 440 completed the survey (15.35% of all candidates trained in 2017 or 2018). Table 3.2 provides a summary of demographic variables for this sample. These candidates had been trained by 63 different providers and those who had been assessed, had been assessed by 47 different providers.

Table 3.2: Participant descriptive statistics.

Sex	n	M_{age}	SD_{age}	White[a]	Assessed[b]	Assessed within 18 months[c]	Passed First Time[d]
Female	155	36.10	10.94	140 (90.32%)	50 (32.3%)	45 (29.03%)	42 (84%)
Male	285	40.77	12.33	263 (92.28%)	119 (41.8%)	108 (37.89%)	103 (86.55%)

[a] Percentage of candidates who are white.

[b] Percentage of candidates who were assessed.

[c] Percentage of candidates who were assessed within 18 months of their training course.

[d] Percentage of candidates who were assessed who passed first time.

When responding to the survey, candidates were at different stages of the pathway. Some candidates had been assessed whereas others had not. In addition, some candidates had completed their training course at least 18 months before responding to the survey, and the remainder responded to the survey less than 18 months after their training course. Therefore, candidates could have completed the survey either prospectively or retrospectively with regards to *both* the event (the assessment itself) and the criterion variable (getting to assessment within 18 months of training) *separately*. As such we were able to create four groups within each sex (see Table 3.3 for descriptive data).

Table 3.3: Candidates pathway progress when completing the survey.

Assessment	18 months post-training	n
Female		
Post-	Greater than or equal to	12
	Less than	22
Pre-	Greater than or equal to	19
	Less than	102
Male		
Post-	Greater than or equal to	36
	Less than	55
Pre-	Greater than or equal to	35
	Less than	159

Each of the three studies had different inclusion criteria and subsequently, used a different subset of candidates. Details of the candidates included in each data set are presented in the sections below and a visual representation of groups that candidates were included in is presented in Figure 3.1.

3.2.2 Measures

We collected data from candidates for this study through two mechanisms. Firstly, we retrieved data from Mountain Training's Candidate Management System (CMS). The CMS data are held by Mountain Training and include information on candidate demographics, training course attendance, and experience data in the form of a digital logbook (DLOG). Secondly, we developed a self-report survey tool to collect quantitative data from Mountain Leader candidates that was not already held by Mountain Training. Given the large number of factors identified as important in Study 1, the first challenge was to create a survey tool that was of reasonable length and would therefore be completed by candidates.

Given the large number of potentially important variables, it was necessary to carry out extensive pilot work, involving two separate studies: one to identify suitable short-form measures of constructs (i.e., one or two items per construct) and another to reduce the number of constructs included in a survey tool, so that we could administer it to candidates without being unduly onerous for them to complete. The resultant

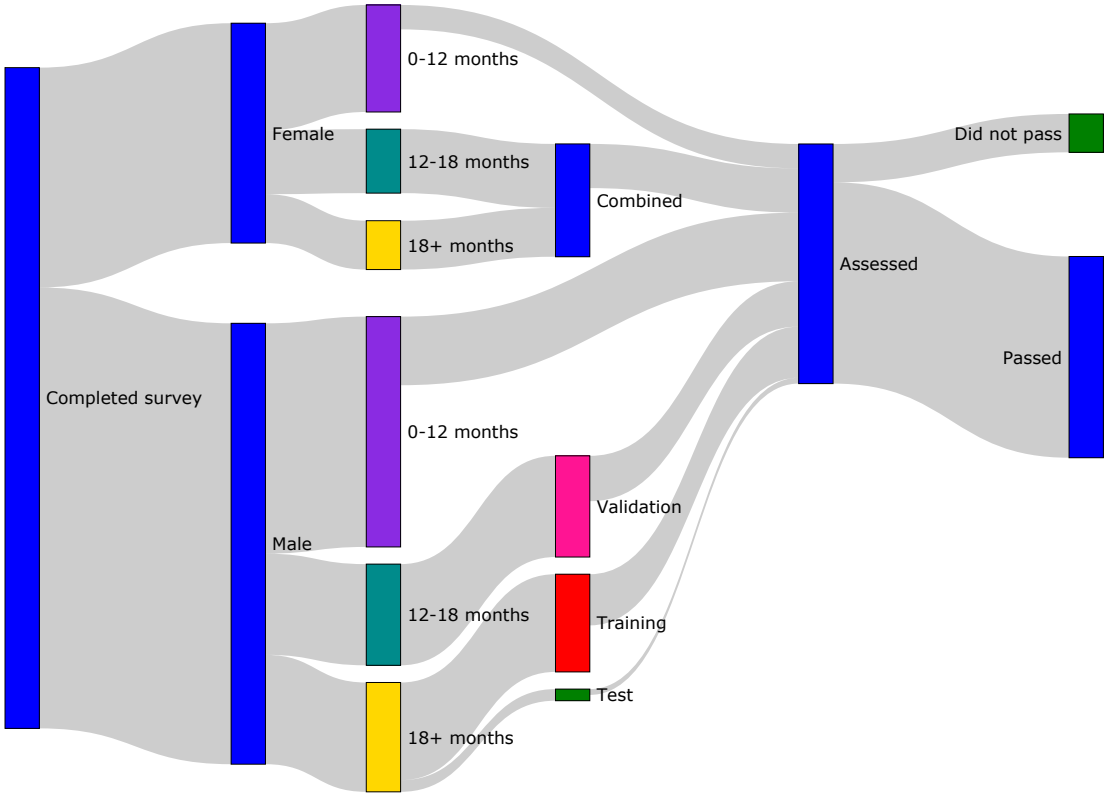

Figure 3.1: Study 4 participants. For simplicity, candidates who have not been assessed have not been added to this figure as a final group, therefore it can be assumed that candidates not progressing from one node to another have not been assessed.

survey tool included constructs within each of the domains listed in Table 3.1. We believe that it is important for the reader to understand the work that underpins the survey tool, although, including that detail here would distract from the purpose of this chapter—which is to understand the factors that best discriminate candidates who are assessed from those who are not. Therefore, the development of this survey tool and a full list of measures included is described in detail in Appendix B. We encourage the reader to engage with this material having read this chapter, to understand the rigour of the pilot work and the techniques employed to reduce the number of items required to measure the constructs of interest, a method which may prove useful in other domains.

3.2.3 Procedure

After the project received institutional ethical approval, we invited Mountain Training candidates who had attended a Mountain Leader training course in 2017 and 2018 to complete the survey tool through the Qualtrics online survey platform (Qualtrics, 2019). Before completing the survey, participants provided informed consent. Following this, they were asked to indicate if they had attended a Mountain Leader assessment course or not, so that they were shown the appropriate questions. We then instructed participants to think about *how they felt before their first assessment* for consolidation and assessment related questions if they had been assessed, and to think *how they felt now* when answering these questions if they had not been assessed when completing the survey.

3.2.4 Analytical Method

We used pattern recognition analyses to identify the most important discriminatory variables within each group. By identifying the most important, we were able to infer which variables were less important discriminatory variables. Pattern recognition analyses, originally developed in bioinformatics (Duda et al., 2000), use machine learning algorithms to identify a set of discriminatory features (variables), which can be used to identify the class (group) of objects (candidates). Pattern recognition analysis is more appropriate for these data than "traditional" methods (e.g., discriminant function analyses) as pattern recognition employs both linear and non-linear functions and therefore reflects multiple and complex interactions and not just "main-effects."

More specifically, we used a pattern recognition procedure that has been developed for analysing what are known as *short and wide* data sets (i.e., data sets that contain more variables than cases) as the present data set are. This pattern recognition procedure has been used in several recent studies to examine differences between athletes of different performance levels (e.g., Güllich et al., 2019; Jones et al., 2019; Jones, 2019; Jones et al., 2020).

This procedure is a three-part process. First, we aimed to identify a set of features which correlated well with the class but had a low correlation with one another (*feature selection*). Second, we tested the ability of this feature subset to correctly classify the candidates according to the criterion variable for that analysis (*classification*). Finally, we refined the feature subset to identify the simplest solution that best explained the data (*recursive feature elimination*). We completed all analyses using WEKA 3-9-3 open source software issued under the GNU General Public License version 3 (Bouckaert et al., 2018; Frank et al., 2016). WEKA is a machine learning workbench with a collection of algorithms widely used for data mining, machine learning, and pattern recognition.

3.2.4.1 Preprocessing.

Using the same data to train and test a model leads to the risk of over-fitting and classification rates being artificially inflated as all the data have been "seen" during the feature selection stage. This phenomenon is known as "peeking" (Kuncheva and Rodríguez, 2018; Reunanen, 2003; Smialowski et al., 2010) and can be avoided by holding some data out of the feature selection stage as outlined below. For a given classification problem, the ideal way to perform the analyses would be as follows:

1. Given N cases, randomly select x cases, where

$$\frac{N}{3} > x > \frac{N}{10}$$

for each class of the criterion variable to be held-out as a test data set D_{test} and the remaining candidates become the training data D_{train}.

2. Prepare both the D_{train} and D_{test} data sets separately (e.g., standardising the data).

3. Perform the feature selection process using D_{train}.
4. Carry out the classification process using D_{train} using k-fold cross validation to select the best model.
5. Carry out the classification process on the previously unseen data (D_{test}) using D_{train} to train the classification model chosen in step 4.

We could only use this full procedure for the data collected from male candidates as there were not enough data available from female candidate or candidates who had not passed their first assessment to. For female candidates and the first-time pass analysis, we included all cases in D_train. D_{train}.

3.2.4.2 Feature Selection.

Feature selection aims to remove irrelevant and redundant variables from the analysis to improve the predictive performance of models (Guyon and Elisseeff, 2003). In this study, we used three techniques designed to improve the performance of feature selection when applied to short and wide data: the use of multiple feature selection algorithms, carrying out feature selection in a *vertically distributed* fashion, and using *leave-one-out cross-validation*.

When using multiple feature selection algorithms, the greater the number of algorithms that select a feature, the more confident one can be that the feature is important as it is less likely that the feature has been chosen by chance (Visa et al., 2011). In this study, we used four feature selection algorithms: Fast Based Correlation Filter (FCBF; Yu and Liu, 2003), Correlation Attribute Evaluator (CAE; Bouckaert et al., 2018), Relief-f (Kira and Rendell, 1992), and Support Vector Machine - Recursive Feature Elimination (SVM-RFE; Guyon et al., 2002).

CAE, Relief-f, and SVM-RFE rank all features in order of merit (magnitude of relationship), whereas FCBF selects a subset of features that are highly correlated with the class but not with one another. As only FCBF provides a subset of features, we selected the top 20 features from the rankings provided by the other three algorithms (if the attribute merit was greater than zero).[2] All four algorithms are well established feature selection methods and the most important point to note about these four

[2]If there were fewer than 20 features in the subset that feature selection was being applied to, we selected the top 10 features. There were more than 10 features in all subsets.

algorithms is that each works in very a different way (see Bolón-Canedo et al., 2015b).

When applied to a data set, using multiple feature selection algorithms yields several feature subsets for each classification problem based on the agreement between the algorithms about the importance of each feature. Features that are not selected or are only selected by one feature selection algorithm are discarded. The following feature subsets are then created from the remaining features: features selected by at least two feature selection algorithms (2s) and features selected by at least three feature selection algorithms (3s). In the studies reported in this chapter we only retained feature subsets which contained at least five features. In this chapter, the feature subsets of features selected by all four feature selection algorithms contained a maximum of four features, therefore they were not retained.

We ran each algorithm using a leave-one-out cross-validation (LOO-CV) protocol. LOO-CV is a special case of K-fold cross-validation, where $K = N$, as it reduces the impact of each object on the feature selection process by increasing the generalisability of the model (Hastie et al., 2009; de Rooij and Weeda, 2020). Each data set was split into K parts or *folds*, with each fold having an approximately equal number of cases. The Kth fold is then removed from the data and the feature selection algorithm is then applied to the remaining data, with each feature being assigned a merit score (or being selected/not for CFS), once this has been repeated K times the merit score for each attribute is averaged across the K iterations.

Feature selection was carried out separately in a both a vertically distributed and *centralised* fashion. Centralised feature selection includes all features at once, whereas vertically distributed feature selection applies the algorithm to several distinct subsets of features, before merging the features selected in each vertical partition, to form a previously unseen feature subset, and applying the algorithm to the new merged feature subset (see Bolón-Canedo et al., 2015b). There is evidence that this process can improve classification rates as it results in "a more balanced feature/sample ratio" reducing the likelihood of overfitting problems (Bolón-Canedo et al., 2015a, p 137). For the getting to assessment analyses the survey data were split into three sections: psychosocial, training, consolidation; and the DLOG data were split into four sections based on time post-training: DLOG experience at training (DLOG_t), DLOG experience six months post-training (DLOG_t6), DLOG experience 12 months post-training (DLOG_t12),

DLOG experience 18 months post-training (DLOG_t18). For the passing first time analyses the survey data were split into the same three sections as for the getting to assessment analyses and the DLOG data were split into three sections: previous course experience, DLOG experience at training (DLOG_t), and DLOG experience at assessment (DLOG_a).

In this study, the combination of multiple feature selection algorithms and vertically distributed meant that we created 2s, and 3s for each feature subset. We then merged the vertical partitions based on level of agreement before reapplying the four feature selection algorithms to the new merged feature subsets. For example, all of the 2s across the feature subsets were combined to form a new merged subset and the feature selection process was then reapplied to this new feature subset, potentially resulting in a further two feature subsets for each classification problem (i.e., merged 2s and merged 3s).

The merging process was carried out for all feature subsets as well as for the survey-based feature subsets and DLOG based feature subsets separately. This process resulted in several candidate feature subsets to be carried forward to the classification stage of the analysis. For each classification problem, there were subsets of features selected by at least two feature selection algorithms (2s) and features selected by at least two feature selection algorithms (3s) for the following candidate feature subsets: centralised, each vertical partition of the data, merged, merged survey, merged DLOG.

3.2.4.3 Classification.

In order to evaluate the predictive performance of each candidate feature subset, we performed *initial classification* experiments using WEKA's Experiment Environment (Bouckaert et al., 2018; Frank et al., 2016). As in the feature selection step, classification experiments used four classification algorithms and LOO-CV given the nature of the data. We used the following classification algorithms with their default settings: Naïve Bayes (NB; John and Langley, 1995), Sequential Minimal Optimization (SMO; Platt, 1998), Instance Based Learning (IBk; Aha et al., 1991), J48 Decision Tree (J48; Quinlan, 1993). As with feature selection, the more consistent the results from each algorithm (classification accuracy) for a feature subset, the more confidence we can place in the predictive validity of that subset. This process returned a classification rate for

each feature subset and classifier.[3] As with the approach taken by Güllich et al. (2019), we then rated each model as *excellent*, *very good*, *good*, *modest*, and *poor* based on our interpretation of the quality of discrimination based on the different percent accuracies.

Having conducted the initial classification experiments, we sought to identify more parsimonious models, potentially with higher classification accuracies using the Recursive Feature Elimination (RFE) method (Guyon et al., 2002). This process is known as *final classification*. To complete final classification, we took each feature subset with more than five features in, examined the normalised SMO weight provided by the SMO classifier and removed the feature with the lowest weight before re-running the four classifiers on the, now smaller, feature subset. This process continued in an iterative fashion until all features with a SMO weight $< .4$ had been removed, the iteration with the best classification rate was then retained as a new feature subset.

3.2.4.4 Model Selection.

The feature selection process yielded 33 getting to assessment models for both male and female candidates and 26 first time pass models. For each of the classification problems listed above, we selected the "best" models. It is important to recognise that these best models are not the only useful ones; however, they were the models that best classified the training data. It is also important to note that we considered the classification profile for each model, rather than just the mean score. It is not uncommon for one classifier—often J48—to perform much worse than the others, therefore if a model performed well with three classifiers and poorly with another, that model was preferred to one that performed better on average (i.e., had a greater mean classification accuracy). For example, consider the classification profiles of the following models, Model A: NB = 85, SMO = 90, IBk = 85, J48 = 50 (mean = 77.5) and Model B: NB = 80, SMO = 80, IBk = 80, J48 = 80 (mean = 80). In this example we would prefer Model A to Model B.

[3]It is important to note that as all of the data have been seen during the feature selection stage the classification rates may be slightly higher than they would be for previously unseen data (Kuncheva and Rodríguez, 2018; Smialowski et al., 2010).

3.3 Study 2: Male Candidates Getting to Assessment

In Study 2 we sought to identify variables that could discriminate male candidates who were assessed within 18 months of their training course from male candidates who were not assessed within 18 months of their training course.

3.3.1 Method

3.3.1.1 Participants.

There were 0 responses from male candidates who completed the survey more than 18 months after their training course (i.e., retrospectively), 0 of whom had been assessed within 18 months of their training course and 0 who not been assessed at the time of completing the survey. A further participants who had been assessed more than 18 months post-training were excluded from the analyses as the wording of the questions they answered meant that their responses were not comparable to those who had been assessed within 18 months and those who had not been assessed. Therefore, from the 65 eligible candidates, we were able to create a set of *training data* ($n = 55$), which we could use to select variables and a set of *test data* ($n = 10$), with five candidates who had been assessed 18 months after their training course and five who had not been assessed 18 months after their training course.

3.3.1.2 Analytical Method.

3.3.1.2.1 Model Testing and Validation. A further 60 male candidates completed the survey less than 18 months after their training but as of 9th June 2020, were at least 18 months post-training. These candidates formed the *validation data* set. We used the test and validation data sets to test the generalisability of models retained from the model selection step of the analysis. To do so, we used the same four classifiers as in the model selection step (i.e., NB, SMO, IBk, and J48) and rather than using LOO-CV, we used the training data to train the classifiers to predict the class of each object in the test and validation data sets. The performance of these predictions was assessed using percentage classification accuracy as in the model selection stage of the analysis.

3.3.2 Results

3.3.2.1 Model Selection.

Using the feature selection method outlined above in Section 3.2.4.2, as noted we created 33 different feature subsets using the training data for classifying male candidates as assessed or not assessed within 18 months of their training course. To evaluate the performance of each feature subset, we carried out initial classification on all 33 feature subsets. There were 13 feature subsets that were *very good* at classifying the data. One of these was from an original feature subset, two were from the centralised feature selection, and the remainder were merged feature subsets from the decentralised feature selection. The performance of each feature subset can be seen in Appendix C Table C.1. We retained the best two of these feature subsets: the "Merged survey 2s 2s" and the "Centralised 3s" feature subsets to carry forward to the final classification step of the analysis. There were 18 unique features between the two feature subsets; six features were common to both feature subsets, and twelve features were contained in only one of the feature subsets.

For the final classification step, we carried out the recursive feature elimination process on the two feature subsets separately, Table 3.4 shows the results of this process. In the Merged survey 2s 2s feature subset only one feature was removed; IBk and J48 saw improvements in classification rates, SMO decreased in performance, and NB remained the same. The RFE feature subset was retained as it had fewer features and had a better classification profile—both in terms of average and consistency. In the Centralised 3s feature subset, again only one feature was removed, improving the performance of NB and SMO, but substantially reducing the performance of IBk, and the performance of J48 remained the same. This time, we retained the original feature subset as it had a better classification profile than the RFE feature subset. Given that neither the Centralised 3s or Merged survey 2s 2s RFE feature subsets performed better than the other, we retained both as predictive models for the model testing and validation steps.

Table 3.4: Male candidates getting to assessment within 18 months of training, classification rates for feature subsets included in final classification.

		Classification rate (%)			
Feature subset	$n_{features}$	NB	SMO	IBk	J48
Initial classification					
Merged survey 2s 2s	17	90.91	92.73	80.00	89.09
Centralised 3s	7	85.45	89.09	90.91	89.09
Final classification					
Merged survey 2s 2s RFE	16	90.91	90.91	89.09	90.91
Centralised 3s RFE	6	89.09	90.91	81.82	89.09

Note: NB = Naïve Bayes, SMO = Sequential Minimal Optimization, IBk = Instance Based Classified, J48 = J48 Decision Tree.

3.3.2.2 Model Testing and Validation.

We tested both models selected above on the test data. Across all four classification algorithms, each model classified the test data with 90% accuracy, except for IBk in the Merged survey 2s 2s RFE model, which had a classification rate of 80%.

Case 8 was misclassified by NB, SMO, and IBk and Case 3 was misclassified by J48 in both models. Case 7 was also misclassified in the Merged survey 2s 2s RFE model by IBk. The performance of these models on the test data is evidence that the models are not over-fitted to the training data as they are similar to the classification rates from the training data, thus increasing our confidence that the variables selected are important discriminatory variables.

When applied to the validation data (i.e., candidates who completed the survey 12-18 months post-training), the performance of the models was *good*, consistent across both the classifiers and models, but was lower than in the test data (see Table 3.5). It is important to note that validation data set included candidates who were assessed more than 18 months-post training, which neither the training nor test data sets did. Therefore, this reduction in classification accuracy may not be surprising.

Table 3.5: Group 5 male candidates getting to assessment within 18 months of training, test and validation data model performance.

	Classification rate (%)			
Feature subset	NB	SMO	IB6	J48
Test data (n = 10)				
Merged survey 2s 2s RFE	90.00	90.00	80.00	90.00
Centralised 3s	90.00	90.00	90.00	90.00
Validation data (n = 60)				
Merged survey 2s 2s RFE	71.67	76.67	75.00	70.00
Centralised 3s	73.33	76.67	76.67	66.67

Note: n = number of candidates, NB = Naïve Bayes, SMO = Sequential Minimal Optimization, IBk = Instance Based Classified, J48 = J48 Decision Tree.

To better understand the prediction errors in the validation data set, we assigned a *voted predicted class* for each candidate based on the average predicted class across the classifier ensemble[4] for both the Merged survey 2s 2s RFE and Centralised 3s models. We then split the candidates into groups based on three factors: when the candidates completed the survey, prospectively (i.e., before an assessment) or retrospectively (i.e., after an assessment); if they had been assessed within 18 months of their training course; and if they had ever been assessed. We then calculated the percentage accuracy of the voted predicted class within the resultant groups and used this to assess the performance of the model, rather than the classification rates of the individual classifiers. The mean classification rates for each group by model combination are shown in Table 3.6.

This analysis shows that, again, both models perform approximately equally well. Both models were extremely good at classifying candidates who had been assessed within 18 months of their training course and responded to the survey after their assessment. Both models were also very good at classifying candidates who were not assessed within 18 months of their training nor had they been at the time of writing ($M_{interval} = 2.65$, $SD_{interval} = 0.1$ years; over 80% of male candidates who are ever assessed are assessed within this time period (see Figure 1.3). These groups are the same as the two groups of candidates who were included in the training and test data sets.

However, the models were less good at classifying candidates who completed the survey prospectively and were subsequently assessed. The models were moderately good

[4]We added a fifth classifier, the Multilayer Perceptron (MLP; Bishop (2006)), for these classification analyses to ensure that there were no ties amongst the predicted classes for a given object.

Table 3.6: Group 5 male candidates getting to assessment within 18 months of training, sub-group prediction model performance.

Survey completion	Assessed within 18 months	Assessed	n	Classification rate (%)	
				Merged survey 2s 2s RFE	Centralised 3s
Prospective	FALSE	FALSE	23	82.61	82.61
		TRUE	10	20.00	30.00
Retrospective	TRUE		27	92.59	88.89

at classifying those who complete the survey prospectively and were assessed within 18 months of their training course and were extremely poor at classifying candidates who completed the survey prospectively and were assessed more than 18 months after their training course. It is important to note that candidates who were assessed more than 18 months post-training were excluded from the training and test data as the wording of the questions shown to them was not comparable to those who had been assessed within 18 months, which was important for questions asking the candidate to consider the "six months before assessment." Further, given the broader context of these analyses—trying to understand the factors that differentiate those who complete the Mountain Leader from those who do not—this poor classification accuracy may not be so important as many of the candidates that are being misclassified are those who are assessed, but more than 18 months post-training.

When candidates who were assessed more than 18 months post-training were excluded from the analysis, the classification rates for both models were 84.89%, which is much closer to the classification rates in the training and test data. These findings provide is further evidence that the models are not over-fitted to the training data and are good at classifying candidates who are assessed within 18 months of their training course.

The results from this data set do allow us to place more confidence in the models selected and many of the classification errors may not be important to Mountain Training as some of the candidates who are being misclassified are still going on to be assessed (most within 24 months of their training).

3.3.2.3 Key Discriminatory Features.

Given that no single model performed better than all others, we retained two classification models for male candidates: a 16 feature model (the Merged survey 2s 2s RFE feature subset) and a seven feature model (the Centralised 3s RFE feature subset). Figure 3.2 shows the normalised group means for the training data of the 16 unique features included in the Merged 2s 2s model and Figure 3.3 shows the normalised group means for the training data of the seven unique features included in the Centralised 3s RFE model. Table 3.7 shows the unstandardised descriptive statistics for the 17 unique variables in the two models for each of the two classes separately. Both representations of the data can be considered as *stereotypical profiles* for the two classes. It is important that the features within the models are considered in a holistic manner as it is their combination that can correctly classify candidates, not any single feature.

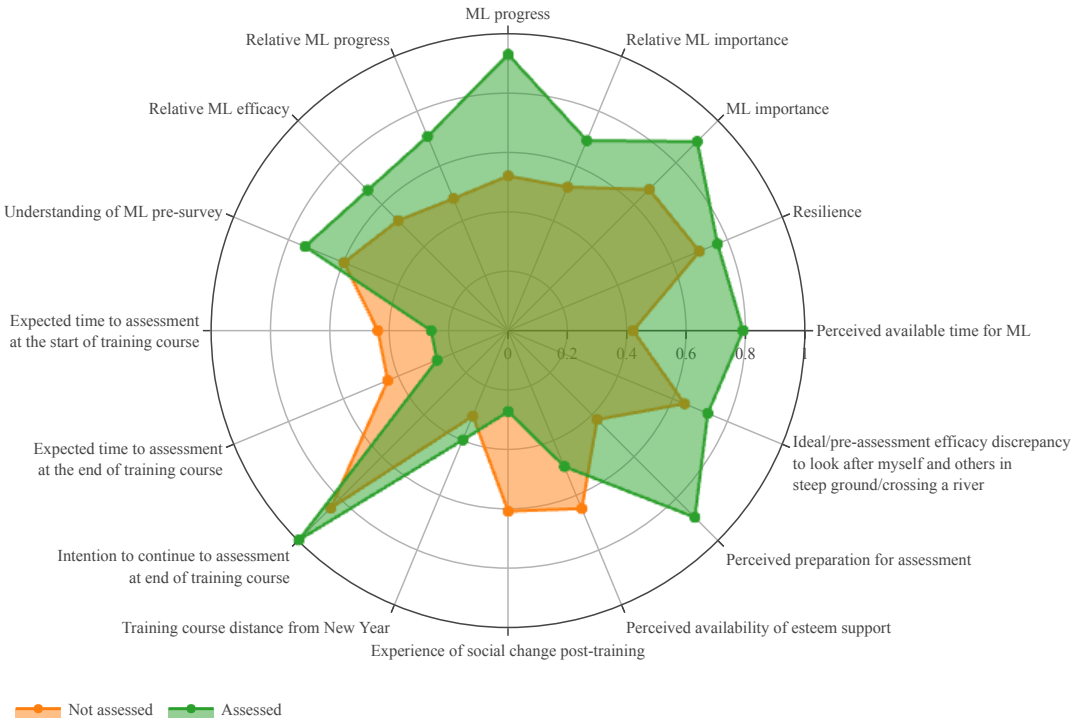

Figure 3.2: Merged survey 2s 2s RFE: Normalised training group means for male candidates getting to assessment within 18 months of their training course.

The analyses for this classification problem identified two predictive models with equal performance; the features included in one of which were a subset of the features included in the other. The models were excellent at classifying both the training and

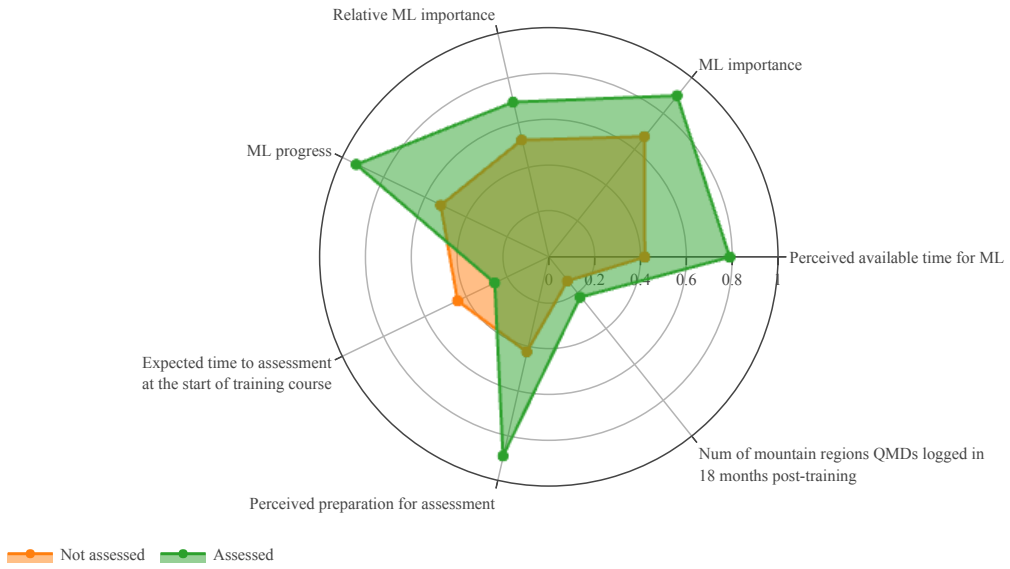

Figure 3.3: Centralised 3s: Normalised training group means for male candidates getting to assessment within 18 months of their training course.

test data, which suggests that the model developed using the training data can be generalised to previously unseen data. The models suggest that, male candidates who are assessed within 18 months of their training course are more likely than male candidates who are not assessed within 18 months of their training course to

- feel that they have enough time to become a Mountain Leader
- feel more resilient
- feel that becoming a Mountain Leader is important
- feel that becoming a Mountain Leader is more important than achieving other life goals
- feel that they have made progress towards becoming a Mountain Leader.
- feel that they have made more progress towards becoming a Mountain Leader than other life goals
- have more confidence in their ability to become a Mountain Leader than to achieve other life goals
- have felt that they had a better understanding of the Mountain Leader qualification before they attended their training course
- have felt that it would take a shorter period of time to get from training to assessment, both at the start of and end of their training course

- have a greater intention to be assessed by the end of their training course.
- be trained in the summer
- experience less social change post-training
- feel that they have less esteem support available
- feel that they had done more to prepare effectively for an assessment in the last six-months
- feel closer to their ideal level of self-efficacy to look after themselves and others in steep ground/crossing a river
- have QMDs logged in a greater number of mountainous regions 18 months post-training

Table 3.7: Unstandardised group descriptive statistics of the features that discriminate male candidates who are assessed within 18 months of their training course from those who are not.

	Not assessed					Assessed				
Variable	mean	median	sd	min	max	mean	median	sd	min	max
Perceived available time to become an ML[ab]	42.07	45.00	27.35	0.0	100.00	79.11	85.00	22.51	13.00	100.00
Resilience[a]	5.33	5.50	1.58	1.5	7.00	5.70	6.00	0.84	3.00	6.50
Importance of becoming an ML[ab]	68.19	69.00	25.57	1.0	101.00	91.00	94.50	11.19	66.00	101.00
Relative importance of becoming an ML[ab]	-21.80	-20.00	25.30	-100.0	17.00	3.54	1.00	14.99	-31.00	49.50
Progress towards becoming an ML[ab]	53.07	59.00	29.69	1.0	90.00	93.96	101.00	13.14	53.00	101.00
Relative progress towards becoming an ML[a]	-13.65	-12.00	33.23	-75.0	44.00	15.30	11.25	17.81	-23.00	52.50
Relative efficacy of becoming an ML[a]	-7.44	0.00	25.61	-74.0	36.00	10.95	5.50	20.79	-27.50	53.00
Recalled understanding of the qualification pre-training[a]	67.52	75.00	23.12	19.0	100.00	78.93	80.00	13.86	41.00	100.00
Expected time to assessment at start of training[ab]	16.93	12.00	9.48	2.0	36.00	10.82	12.00	4.40	3.00	24.00
Expected time to assessment at the end of the training course[a]	16.93	12.00	9.48	2.0	36.00	10.82	12.00	4.40	3.00	24.00
Intention to complete at the end of training[a]	87.89	100.00	22.74	22.0	100.00	99.75	100.00	1.00	95.00	100.00
Training course distance from New Year[a]	-0.38	-0.35	0.44	-1.0	0.37	-0.21	-0.14	0.59	-0.99	0.99
Experience of social change post-training[a]	60.81	70.00	33.86	0.0	100.00	27.25	7.50	33.75	0.00	94.00
Perceived availability of esteem support[a]	3.59	4.00	1.05	1.5	5.00	2.98	3.00	1.12	1.00	5.00
Perceived preparation in the last six months/six-months before assessment[ab]	42.33	32.00	33.69	0.0	100.00	88.89	91.00	13.22	48.00	100.00
Difference between ideal and pre-assessment efficacy to look after myself and others in steep ground/crossing a river[a]	-13.30	-9.00	17.97	-57.0	11.00	-7.54	-5.50	8.20	-25.00	0.00
Number of mountain regions QMDs logged in 18 months post-training[b]	13.44	14.00	12.26	0.0	49.00	22.54	17.00	20.60	0.00	101.00

[a] Included in the merged survey 2s 2s RFE model
[b] Included in the centralised 3s model

3.3.3 Discussion

We identified two predictive models that were both very good at correctly classifying the training data (up to 90.91% accuracy) and test data (up to 90% accuracy) for male candidates getting to assessment within 18 months of training. These models were also able to predict if male candidates who completed the survey more than 12 months, but less than 18 months post-training would be assessed within 18 months of their training with good accuracy (up to 76.67%). The variables selected as important discriminatory variables for male candidates getting to assessment within 18 months of their training course can be considered under two main headings: *Context*, how becoming a Mountain Leader fits into a candidate's life and *Self-Efficacy and Resilience*, how confident a candidate is that they can become a Mountain Leader and perform assessment related tasks and how well they bounce back from setbacks. Below we discuss the variables selected and then the performance of the models on the test and validation data sets. To facilitate discussion of the results, we have grouped the variables into three categories: context, self-efficacy, and other.

3.3.3.1 Features selected.

3.3.3.1.1 Context. Several variables relating to how becoming a Mountain Leader is situated in the wider context of male candidates' lives were important for discriminating male candidates who had been assessed 18 months post training from male candidates who had not been. The attitudes of a candidate towards being assessed will likely be informed by the importance of becoming a Mountain Leader, their understanding of the qualification, and the time they expect it will take them to get from their training course to an assessment. Put another way, how a candidate feels that becoming a Mountain Leader will fit into, or even enable, their life may influence their attitude towards being assessed. In addition, the amount of time that candidates feel they have available to become a Mountain Leader in and their perceived efficacy to become a Mountain Leader will likely influence their perceived behavioural control. The *theory of planned behaviour* (Ajzen, 1991; Ajzen and Madden, 1986) provides a useful framework for discussing many of these variables. According to the theory of planned behaviour the attitudes and perceived behaviour control formed by these variables would form a candidate's intention to be assessed. In addition to this, there is evidence in the personal goal

literature that goal importance influences how committed to a goal an individual will be (Gollwitzer, 1993) and that individuals will engage in task-relevant behaviours when they feel that the task is important (e.g., Ingledew et al., 2005; Yukl et al., 1999, 1996).

Candidates reported their intention to be assessed at various stages of the pathway (at registration, at the start of their training course, and at the end of their training course), but only their intention to be assessed at the end of their training course was selected as an important discriminatory variable. Not all candidates have a good understanding of the Mountain Leader qualification and therefore may be attending a training course to find out more about the qualification (see also Study 1). This understanding may influence candidates' attitudes towards being assessed and it is also likely to influence their intention to be assessed. Therefore, we would suggest that the strength of a candidate's intention to be assessed at the end of the training course is more important than their intention at the start of the training course because it is based on a more accurate and complete understanding of the qualification.

Candidates who had not been assessed when they responded to the survey were also asked to report their intention to be assessed at that point in time ($n = 334$). Analysis of these data (reported in Appendix D) suggest that candidates with a greater intention to be assessed were more likely to be assessed six months after answering the survey than those who reported a lesser intention to be assessed. These results support the hypothesis, from the theory of planned behaviour, that intentions cause behaviours rather than this feature having been selected due to attribution bias.

Whilst the observation that candidates who expect it to take them longer to get from training to assessment were less likely to be assessed within a given period may seem elementary, it is important to note that, of the 27 candidates not assessed within 18 months of their training in the training data, only 10 expected it to take them more than 18 months. One reason for candidates expecting it to take them longer to get from training to assessment may be that they feel they have less available time to become a Mountain Leader than those who expect it to take less time. We were unable to test the direction of this hypothesis with the data from the present study, but it would seem more likely that the less time a candidate feels they have available to become a Mountain Leader, the longer they would expect it to take them to get to an assessment, rather than thinking that they do not have time available to become a Mountain Leader

because they expect it to take them longer to get to assessment.

In addition, there is a substantial body of evidence that suggests that goals that are proximal in time are more likely to be adhered to (see Hardy et al., 1996; Weinberg and Gould, 2014), therefore we would expect candidates who expect to be assessed sooner to be more likely to get to assessment than those who expect it to take a longer time. Further, experiencing social change after a training course may mean that candidates have more or less available time and/or change their importance in becoming a Mountain Leader. The question used in the survey did not ask if candidates had more or less resources (e.g., available time) because of this change, however given that more social change a candidate experienced, the less likely they were to be assessed within 18 months, it would be reasonable to assume that these social changes are more likely to leave candidates with less, rather than more, resources to become Mountain Leaders. If candidates who expect to take longer do take longer, then there will be more opportunities for barriers to prevent them pursuing and/or attaining that goal. Indeed, experiencing social change may be one reason for the observed difference between expected and actual times to assessment from training.

3.3.3.1.2 Self-Efficacy and Resilience. Self-efficacy theory (Bandura, 1977, 1982) may also be useful in understanding the relationships between some of the variables selected as important discriminatory variables (n.b., Ajzen, 1991, suggested that perceived behavioural control is similar to the construct of self-efficacy). The data presented in Table 3.7 show that, on average, male candidates who were assessed within 18 months of their training course felt more confident in their ability to become a Mountain Leader than they did to achieve other personal life goals, whilst those who were not assessed felt less or equally confident, in their ability to become a Mountain Leader in comparison to other personal life goals. As outlined in Table 3.1, higher levels of self-efficacy have been associated with higher levels of goal progress, task engagement, goal commitment for self-set goals, and on-task effort.

Goals that are too difficult or unrealistic may not be accepted (Kyllo and Landers, 1995). It has also been suggested that goals that are too difficult may be perceived as threatening, therefore strong self-efficacy to meet goals is important for sustained motivation (Bueno et al., 2008). This may suggest that if candidates had lower levels of self-efficacy, they have could have felt the goal of becoming a Mountain

Leader was too difficult or unrealistic, consequently, it would be less likely that they had sustained motivations. This may be one explanation for candidates who did get to assessment having higher scores on self-efficacy constructs than candidates who did not get to assessment within 18 months of their training course.

In addition, there is evidence in the literature that self-efficacy and goal importance interact in such a way that when goal importance is low, self-efficacy has a weaker effect on goal progress, including in situations where multiple goals are considered (e.g., Beattie et al., 2015; Kernan and Lord, 1990; Orbell et al., 2001; Schmidt and DeShon, 2007; Schmidt et al., 2009; Schmidt and Dolis, 2009). Beattie et al. (2015) showed that self-efficacy predicted goal progress when goal importance was high and had no effect when goal progress was low. Considering this finding in relation to the present data, we suggest that the more important a candidate feels it is that they become a Mountain Leader, (a) the more committed they will be to achieving it, (b) the more they will engage in task-relevant behaviours (e.g., preparing for an assessment), and (c) that goal progress would be greatest amongst candidates who felt it was both important that they became a Mountain Leader and also felt that they were able to.

This is congruent with the results discussed above (Section 3.3.3.1.1), where candidates who feel that becoming a Mountain Leader fits into their life are more likely to have been able to commit resources towards preparing for a Mountain Leader assessment. We would suggest that the more resources that a candidate has put towards becoming a Mountain Leader, the higher their levels of self-efficacy will be as a result of them having done more to prepare for an assessment. Therefore, how becoming a Mountain Leader fits into a candidates life may have an indirect effect on their self-efficacy through the preparation they have or have not done for a Mountain Leader assessment.

Preparation for an assessment will likely include deliberate practice of the skills required for candidates to look after themselves and others in steep ground and when crossing rivers. For some candidates, these skills will be the most specialist mountaineering skills they possess, and they will have little reason, beyond passing a Mountain Leader assessment, to practice them. Unless these candidates have spent time deliberately preparing for an assessment, it is likely that they will feel less confident than their ideal-self would at assessment that they can successfully demonstrate these skills.

Therefore, these candidates may feel that they are not ready to pass an assessment and therefore not attend one, a phenomenon described by the participants in Study 1.

Self-efficacy theory suggests that previous performance accomplishments, followed by vicarious experiences, verbal persuasion, and emotional arousal have the greatest effect on self-efficacy and a negative experience of a given magnitude will have a greater effect than an equivalent positive experience (Bandura, 1977, 1982). Candidates who have visited different mountainous regions are more likely to have encountered a range of different terrain and situations that require them to practice their skills—not limited to the specialist mountaineering skills previously mentioned. In doing so, they may have had mastery experiences in a range of settings that increased their levels of self-efficacy to become a Mountain Leader and to carry out tasks related to the assessment.

In our analyses we used the time of year that courses took place as a proxy measurement of weather and daylight hours. We would expect that courses nearer to the New Year would have worse weather and less daylight than those nearer to the middle of the year (i.e., June/July). Given that candidates who were trained closer to the middle of the year were more likely to have been assessed 18 months post-training, we would suggest that better weather and more daylight on the training course provides candidates with a more positive experience and possibly a better learning environment. Such environments may be more conducive to mastery experiences, which would build candidates' self-efficacy. To investigate this further, weather data (held on CMS) and daylight hours data should be included in any future studies.

There is a broad literature reporting the benefits of resilience (e.g., Seery and Quinton, 2016; Smith et al., 2008). Becoming a Mountain Leader is a difficult process which requires the investment of time, energy, and money and most candidates will have to deal with setbacks during this process. Candidates who are more resilient will be better able to overcome adversity (Smith et al., 2008). For example, bad weather on a training course or changes in life circumstances that become barriers to becoming a Mountain Leader. It is also a central tenet of self-efficacy theory that people with firmly established self-efficacy beliefs are more resilient (Bandura, 1997) as the stronger self-efficacy beliefs are, the easier they are to maintain following disconfirming events.

3.3.3.1.3 Other. Social support is typically seen as beneficial to performance and self-confidence (Rees and Freeman, 2007). It may therefore be surprising that

candidates' who were not assessed felt that they had more esteem support available to them than the candidates who had been assessed did. One explanation for this is that candidates who do not feel that they need esteem support answer this question in a different way to those who do (i.e. they don't perceive it as available), therefore those who feel they need it score more highly and with less variation in their responses. Another explanation is that esteem support may be reinforcing beliefs around unpreparedness for male candidates, with greater levels of esteem support acting to simply reminding candidates that they are not ready for an assessment. Without further investigation both explanations remain somewhat speculative, although it is worth noting that findings consistent with the latter explanation, where psychological skills and strategies have paradoxical effects on performance, have been reported elsewhere in the literature (Roberts et al., 2013). Regardless, the results highlight that some support strategies might need to be used with caution. Future research could investigate the potential negative effects of social-support to better understand when support strategies are appropriate.

3.3.3.2 Model Testing and Validation.

The results from the test data set demonstrate that the three-step feature selection process applied to the training data resulted in a model that was not over-fitted to the data and can therefore be generalised beyond the training data. Interestingly, the reduction in classification rate one may expect due to the peeking effect (Kuncheva and Rodríguez, 2018; Reunanen, 2003) was not observed in these data. Whilst we recognise that the absence of evidence is not evidence of absence, this is one of few studies that we are aware of which have used this three step process and then tested the predictive models identified on previously unseen data (see Jones, 2019; Jones et al., 2020) and none of these studies have seen a reduction in the classification rates when models are applied to unseen test data. This observation allows us to place more confidence in the three-step method when applied to similar classification problems (i.e., classification problems using multifaceted, short and wide data sets).

There are three main reasons the classification rates for the models when applied to the validation data may be lower than in the training and test data. Firstly, in the training and test data candidates who had been assessed more than 18 months

post-training were excluded from the analyses. Secondly, we collected the validation data from candidates who had completed their training less than 18 months prior to training; these candidates were not at the same point in the pathway when they answered the survey, as those included in the training and test data were. Therefore, their answers may have been different to those they would have given six-months later, which would be particularly important for variables that asked candidates about the last six-months. For example, a candidate may not have felt that they had made much progress towards becoming a Mountain Leader 6-12 months post-training and then may make progress in the 12-18 month period. Finally, it is possible that candidates are creating narratives in their minds based on whether they have been assessed and that had they answered the survey before being assessed, then their answers may have been different.

The performance of the two models on the training, test, and validation data sets allows us to place confidence in the importance of the features included in the models for discriminating male candidates who are assessed within 18 months of their training course from those who are not.

3.4 Study 3: Female Candidates Getting to Assessment

3.4.1 Method

3.4.1.1 Participants.

We collected the data for this analysis from 27 candidates who had been assessed 18-months after their training and 27 who had not. We received fewer responses from female candidates and were therefore unable to have separate test or validation data sets. In each group there were 10 candidates who completed the survey retrospectively (i.e., more than 18 months post-training) and 17 who completed the survey prospectively but had completed their training at least 12 months before (i.e., 12-18 months post-training).

3.4.2 Results

Using the same feature selection method as for male candidates, we created 37 feature subsets using the training data, which we then used to perform the initial classification procedure on. As in the male candidates' data, there was no single standout feature subset, in this instance two feature subsets (Merged survey 3s 2s and Merged 3s 3s) classified the training data approximately equally well. These two feature subsets were "very good" at classifying the training data and contained 31 unique features, 5 features were included in two of the feature subsets. Table C.2 shows the classification rates for all feature subsets included in the initial classification step.

Both feature subsets were carried forward to the final classification step of the analysis. Seven features were removed from the Merged survey 3s 3s feature subset, improving the classification rate of all four classifiers and three features were removed from the Merged 3s 3s feature subset, however, this process did not improve the performance of any classifier, rather, it reduced the performance of two of them (see Table 3.8). The Merged survey 3s 2s RFE model (Figure 3.4) and the Merged 3s 3s (Figure 3.5) feature subsets were retained as predictive models. Table 3.9 shows the unstandardised descriptive statistics for the 22 unique variables in the two models, for each of the two classes separately.

Table 3.8: Female candidates getting to assessment within 18 months of training, classification rates for feature subsets included in final classification.

		Classification rate (%)			
Feature subset	$n_{features}$	NB	SMO	IBk	J48
Initial classification					
Merged survey 3s 2s	23	87.50	92.86	85.71	78.57
Merged 3s 3s	14	92.86	92.86	92.86	94.64
Final classification					
Merged survey 3s 2s RFE	14	94.64	94.64	85.71	80.36
Merged 3s 3s RFE	11	91.07	92.86	91.07	94.64

Note: NB = Naïve Bayes, SMO = Sequential Minimal Optimization, IBk = Instance Based Classified, J48 = J48 Decision Tree.

The two models retained include 22 features; six features are included in both models, and 16 features are unique to a single model (nine of which were DLOG variables, therefore could only be included in one model). This level of overlap between the models allows us to place more confidence in the importance of the features included

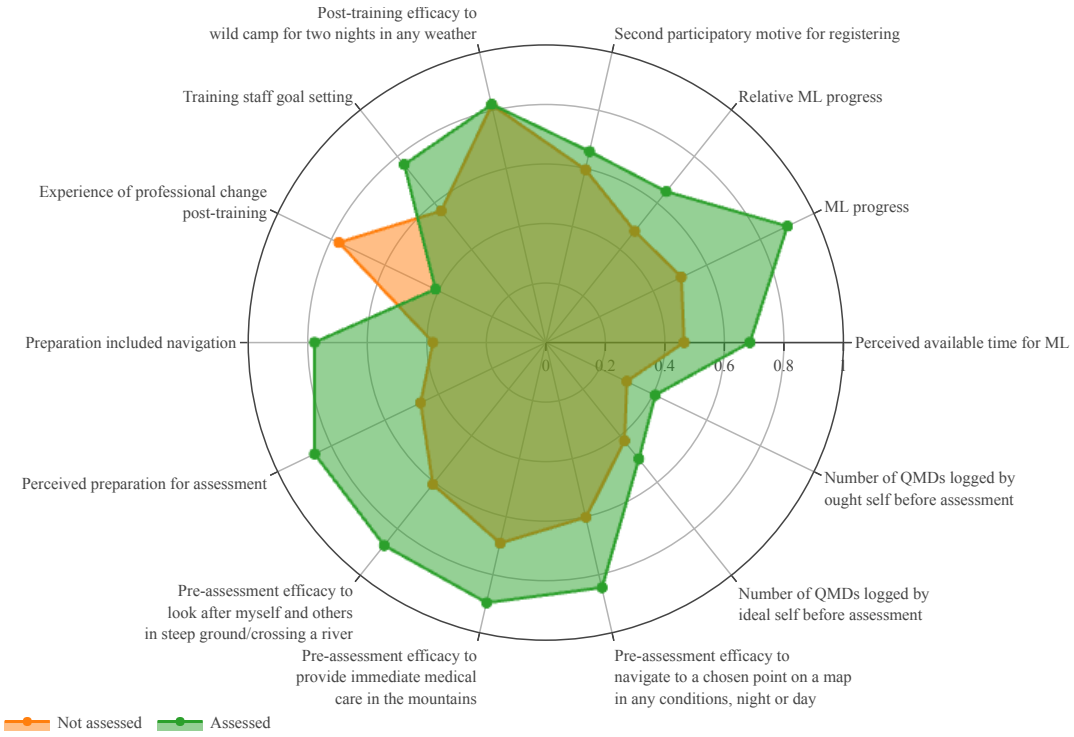

Figure 3.4: Merged survey 3s 2s: Normalised training group means for female candidates getting to assessment within 18 months of their training course.

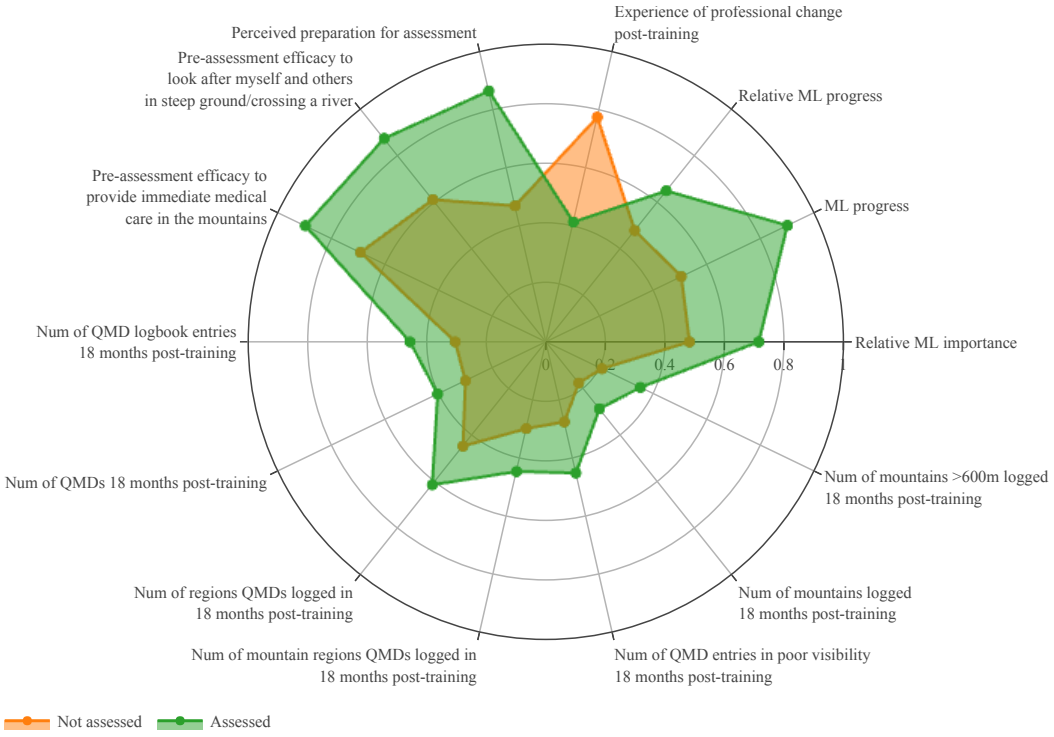

Figure 3.5: Merged 3s 3s: Normalised training group means for female candidates getting to assessment within 18 months of their training course.

in them. The features contained in the two models retained suggest that female candidates who are assessed within 18 months of their training course are more likely than female candidates who are not assessed within 18 months of their training course to

- feel that they have enough time to become a Mountain Leader
- feel that becoming a Mountain Leader is important
- feel that becoming a Mountain Leader is more important than achieving other life goals
- feel that they have made progress towards becoming a Mountain Leader
- feel that they have made more progress towards becoming a Mountain Leader than other life goals
- feel that their training course staff helped them to set goals
- feel more confident at the end of training that they could wild camp for two nights in any weather
- experience professional change post-training
- have an extrinsic second participatory motive for registering
- feel that they have done more to prepare for an assessment in the last six-months
- have practiced their navigation skills
- feel more confident pre-assessment in their skills to
  - look after themselves and others in steep ground
  - provide immediate medical care in the mountains
  - respond appropriately to an emergency
  - navigate to a chosen point on a map in any conditions, night or day
- feel that both their ideal and ought selves would have logged a greater number of QMDs pre-assessment
- eighteen months post-training:
  - have a greater number of QMD entries
  - have a greater number of QMDs[5]
  - have visited more mountainous regions to log their QMDs
  - more QMDs in poor visibility
  - summited more mountains

[5]QMD entries can span more than one day.

– summited more mountains at least 600m high

Table 3.9: Group 5 female candidates getting to assessment within 18 months of training, unstandardised group descriptive statistics

	Not assessed			Assessed		
Variable	mean	median	sd	mean	median	sd
Perceived available time to become an ML[a]	44.37	50.0	29.32	68.44	74	22.03
Progress towards becoming an ML[ab]	49.04	58.0	22.74	90.96	101	15.09
Relative progress towards becoming an ML[ab]	-15.52	-15.5	29.46	17.96	8	26.46
Relative importance of becoming an ML[b]	-29.96	-27.5	29.62	3.43	0	15.39
Second participatory motive for registering[a]	3.04	2.0	1.26	3.63	4	1.28
Post-training efficacy to wild camp for two nights in any weather[a]	90.85	98.0	13.16	85.41	100	20.66
Perception of training course staffs' "goal setting" skills[a]	6.26	6.0	2.54	8.11	9	2.52
Experience of professional change post-training[ab]	83.89	94.0	26.27	41.22	29	35.14
Preparation included navigation[a]	0.41	0.0	0.50	0.78	1	0.42
Perceived preparation in the last six months/six-months before assessment[ab]	43.19	50.0	31.84	86.30	91	17.36
Pre-assessment efficacy to...						
look after myself and others in steep ground/crossing a river[ab]	66.19	70.0	24.54	88.37	91	11.01
provide immediate medical care in the mountains[ab]	73.63	78.0	25.04	90.52	91	8.03
navigate to a chosen point on a map in any conditions, night or day[a]	63.70	67.0	25.62	85.85	90	14.87
Number of...						
QMDs logged by ideal self before assessment[a]	50.56	50.0	13.03	60.00	60	14.14
QMDs logged by ought self before assessment[a]	42.41	40.0	8.70	52.59	50	16.49
QMD logbook entries 18 months post-training[b]	13.07	13.0	7.63	21.48	21	7.89
QMDs 18 months post-training[b]	16.04	16.0	11.25	22.63	21	8.18
regions QMDs logged in 18 months post-training[b]	6.52	7.0	3.45	9.81	10	3.69
mountain regions QMDs logged in 18 months post-training[b]	12.56	13.0	7.44	20.07	20	8.82
QMD entries in poor visibility 18 months post-training[b]	4.19	3.0	3.73	7.67	7	4.39
mountains logged 18 months post-training[b]	30.93	26.0	20.15	51.11	47	32.79
mountains >600m logged 18 months post-training[b]	27.56	24.0	18.72	48.04	42	30.89

[a] Included in Merged survey 3s 2s RFE

[b] Included in Merged 3s 3s

3.4.3 Discussion

There were two predictive models retained that could correctly classify female candidates as having been assessed or not within 18 months of their training course with very good accuracy (up to 94.64%). There were 22 unique variables between the models, six of which were also selected in the final models for classifying male candidates.

Some of the variables discussed under the context heading for male candidates were also selected as important variables for discriminating female candidates who are assessed within 18 months of their training course from those who were not. The relationships discussed above for male candidates between perceived available time to become a Mountain Leader, perceived progress towards becoming a Mountain Leader (both absolute and relative to other life goals), importance of becoming a Mountain Leader (both absolute and relative to other life goals), and progress towards becoming a Mountain leader relative to other life goals and getting to assessment within 18 months of training are likely to be the same for female candidates. However, there are other variables, from various stages of the pathway, that are also important for discriminating female candidates who are assessed 18 months post-training from those who are not. These additional variables can be considered under two headings: motivation and consolidation of experience (i.e., what a candidate does after their training course to consolidate what they have learnt on their training course and prepare for an assessment).

3.4.3.1 Motivation.

In this study, we asked candidates to state two reasons that they had registered for the Mountain Leader qualification (i.e., two participatory motives). For their first reason, there was little variation and most candidates said that they had registered in order to become a Mountain Leader (n.b., this is an extrinsic participatory motive). Female candidates who gave an extrinsic participatory motive for their second motive (e.g., "Pass assessment.") rather than a more intrinsic one (e.g., "To spend more time in the mountains.") were more likely to have been assessed 18 months after their training course. This finding suggests that having more than one extrinsic participatory motive is important for candidates getting to assessment.

This finding does not fit comfortably with the traditional view that

predominantly intrinsic motives are better predictors of performance and goal persistence than predominantly extrinsic motives. However, there is evidence from the sporting domain that self-determined forms of motivation are not always the best predictors of a desired outcome (e.g., Chantal et al., 1996; Fortier et al., 1995; Güllich et al., 2019; Hardy et al., 2017). Indeed, the pursuit of an Olympic gold medal could be likened to the attainment of a qualification, in that ultimately attaining either goal is contingent on external reward and one may argue that it is therefore not possible to be truly intrinsically motivated to attain such goals.

3.4.3.2 Consolidation of experience.

Goal setting facilitated by course staff was also selected as an important discriminatory variable for female candidates. Goal setting has been shown to improve outcomes in several domains (see Weinberg and Gould, 2014, p 356). One way that goal setting facilitated by training course staff may have helped candidates is by enabling them to maximise the benefits of the time that they spent consolidating their skills and preparing for a Mountain Leader assessment after the training course. In addition to this, goal setting may have made it more likely that candidates would prepare for an assessment. The more specific these goals are, the more they will have focused candidates' attention and efforts towards being at the right level to pass an assessment. Further, goal setting will have helped facilitate mastery experiences thus, this goal setting will have helped female candidates develop their confidence, which as discussed below, is key for female candidates getting to assessment.

Female candidates who experienced professional change post-training were less likely to be assessed 18 months post-training than those who did not. For male candidates, we suggested that experiencing social change post-training may reduce the time that they have available to become a Mountain Leader and the same point may also be true for female candidates who experience professional change. An additional consideration for those who experience professional change is that if they were pursuing the Mountain Leader qualification for work, they may no longer need to become a Mountain Leader because of work changes. When considered with the apparent importance of having extrinsic participatory motivation—often for work—Whilst this hypothesis is somewhat speculative given the lack of data to test it with, this

explanation seems plausible and would be characterised by reduced levels of extrinsic motivation and increased levels of amotivation.

Female candidates who were assessed 18 months post-training felt that they had done more to prepare for an assessment in the six-months prior to their assessment than those who had not been assessed felt that they had done in the six-months prior to completing the survey. In addition, they were more likely to have included specific navigation practice in their preparation. Self-efficacy theory suggests that those who have done more to prepare would be more confident in their skills. Whilst the relationship between confidence and experience may vary between individuals (Weinberg and Gould, 2014), mastery experiences will enhance one's confidence (Bandura, 1977) and, in general, the more experience one has the more confident and competent they will be. Indeed, variables related to three different sets of skills were identified as important for discriminating female candidates who did and did not pass their first assessment: navigation, security in steep ground, and emergency procedures. It is important for candidates to be competent in all three of these areas to pass an assessment.

In the Merged survey 3s 2s model (that did not include DLOG variables), the number of QMDs that a candidate felt their ideal- and ought-selves (i.e., their self-guides) would have logged before assessment were included as important discriminatory variables. The greater the number of QMDs logged by a self-guide at assessment, the greater the discrepancy post-training. According to self-discrepancy theory, those with greater discrepancies would be more motivated than those with smaller discrepancies (Higgins, 1987) and they may therefore be more likely to prepare for an assessment. It is also possible that the more diffcult goals (as represented by the self-guides) of the candidates who were assessed within 18 months of their training course were a result of them having higher levels of self-efficacy (cf. Bandura, 1986).

Due to the cross-sectional design of this study, it is not possible to empirically test the direction of causality between motives, confidence, and experience. However, we asked candidates to think about themselves at different points in time (i.e., at registration, training, and pre-assessment) whilst completing the survey and to answer the questions in relation to how they remember feeling at those time points. Assuming that they did this and that attribution bias did not overly influence their answers, we would suggest that their goals (as represented by their self-guides) influenced their

motivation to gain experience, as candidates who set themselves higher goals would have greater levels of motivation to reduce the discrepancy between their actual-self and their self-guides. If this greater motivation resulted in candidates gaining more experience as they prepared for an assessment, they would likely have more mastery experiences, resulting in higher levels of self-efficacy.

We conducted an additional exploratory analysis (described below) to better understand the relationship between self-guides and getting to assessment within 18 months of training. This analysis revealed that only female candidates who were not assessed within 18 months of their training course felt that their ideal- or ought-self would have logged fewer than 40 QMDs pre-assessment—the prerequisite number of QMDs that a candidate must have before passing an assessment (Mountain Training UK, 2015a)—and that all female candidates who were assessed within 18 months felt that their self-guides would have at least 40 QMDs prior to being assessed. In addition, female candidates who felt both of their self-guides would have logged more than 40 QMDs pre-assessment were more likely to have been assessed 18 months post training than those who did not feel that both of their self-guides would have more than 40 QMDs at assessment. The thresholds of 40 QMDs were identified by applying a J48 classifier (with LOO-CV) to the ideal- and ought-self number of QMDs pre-assessment, which had a classification rate of 72.22% and a balanced confusion matrix. When assessed using a χ^2-test, those who exceed these thresholds are significantly more likely to have been assessed 18 months post-training that those who do not, $\chi^2(1, n = 54) = 12.54$, $p < .001$.

In the model developed from the full set of features (i.e., including survey and DLOG variables), neither of the self-guide features, nor the discrepancy scores between the self-guides and the number of QMDs logged 18 months post-training were selected as important discriminatory variables. However, several DLOG variables were selected; this difference in model structure suggests that actual experience is more important than goals for discriminating female candidates who have been assessed 18 months post-training from those who have not. The DLOG variables selected were all related to QMDs and no other types of experience. Candidates who had been assessed 18 months post-training had more logged experience (as represented by the number of QMDs and QMD entries), had gained that experience in a greater range of locations (number of

areas and number of mountain regions), had logged more experience in poor visibility, and had summited a greater number of mountains (in total and mountains higher than 600m above sea level). This set of features suggest that it is not only important that female candidates gain experience during their preparation for assessment, but this experience should be varied and include experiences where their skills will be challenged (e.g., for example having to navigate due to poor visibility), which will provide mastery experiences that will increase their self-efficacy. Whilst the exact nature of the relationship between the features relating to candidates' navigation skills/experience and the outcome is unclear, when considering self-efficacy theory and the results from Study 1, it is likely that this deliberate practice increases female candidates' efficacy to become MLs.

3.5 Study 4: Passing First Time

3.5.1 Method

3.5.1.1 Participants.

We collected the data used for this study from 46 candidates, 34 of whom had been assessed prior to completing the survey and 12 of whom had not been assessed before completing the survey. As with the data in female candidates getting to assessment, we combined the retrospective and prospective data to increase the sample size.[6] Twenty three of the 46 candidates passed their assessment first time. Of the 23 who did not pass, 6 completed the survey prospectively. Two of the 23 candidates who did not pass withdrew from their first assessment, none failed, and the remainder were deferred. Seven of those who were deferred only needed to log additional days.

3.5.2 Results

The feature selection process yielded 26 feature subsets that were all carried forward to the initial classification step. Again, there was no single feature subset that out-performed the others. The two best feature subsets (Merged survey 2s 3s and

[6]This means that no variables about candidates' experience of assessment were included in the analyses. We have run the analyses on just the retrospective data, which allowed us to include some variables about candidates' experiences of assessment, but none of these variables were selected in the best discriminatory subsets, nor were the classification rates significantly higher.

Centralised 2s) were "good" at classifying the data. These two feature subsets contained 20 unique variables, six of these features are included in both feature subsets. Table C.3 shows the classification rates for all feature subsets included in the initial classification step.

These two feature subsets were carried forward to the final classification step. Two features were removed from the Merged survey 2s 3s feature subset and six features were removed from the Centralised 2s model. In both instances, this process resulted in the classification rates for two classifiers improving and one decreasing; these changes made the classification profile more consistent (see Table 3.10). On this basis the Merged survey 2s 3s RFE model (Figure 3.6) and the Centralised 2s (Figure 3.7) were retained as predictive models. Table 3.11 shows the unstandardised descriptive statistics for each feature, retained in the two models, within the two classes.

Table 3.10: Passing first time, classification rates for feature subsets included in final classification.

		Classification rate (%)			
Feature subset	$n_{features}$	NB	SMO	IBk	J48
Initial classification					
Merged survey 2s 3s	8	73.91	67.39	82.61	60.87
Centralised 2s	18	76.09	69.57	82.61	60.87
Final classification					
Merged survey 2s 3s RFE	6	73.91	69.57	71.74	63.04
Centralised 2s RFE	12	82.61	78.26	73.91	60.87

Note: NB = Naïve Bayes, SMO = Sequential Minimal Optimization, IBk = Instance Based Classified, J48 = J48 Decision Tree.

The two models retained include 14 features; four features are included in both models, and ten features are unique to a single model (eight of which were DLOG features, therefore could only be included in one of the models). The features in the two classification models suggest that candidates who pass their first assessment are more likely than those who do not to

- live closer to the mountains
- feel that the training course staff were involved in their development
- experience less inter-personal conflict on their training course
- feel that their course staff used observation and effective questioning skills more often

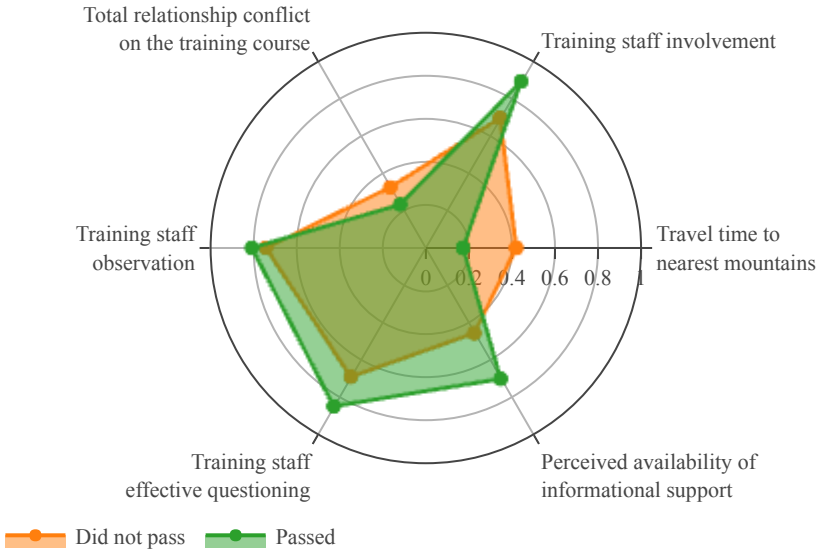

Figure 3.6: Merged survey 2s 3s: Normalised training group means for candidates passing their first assessment.

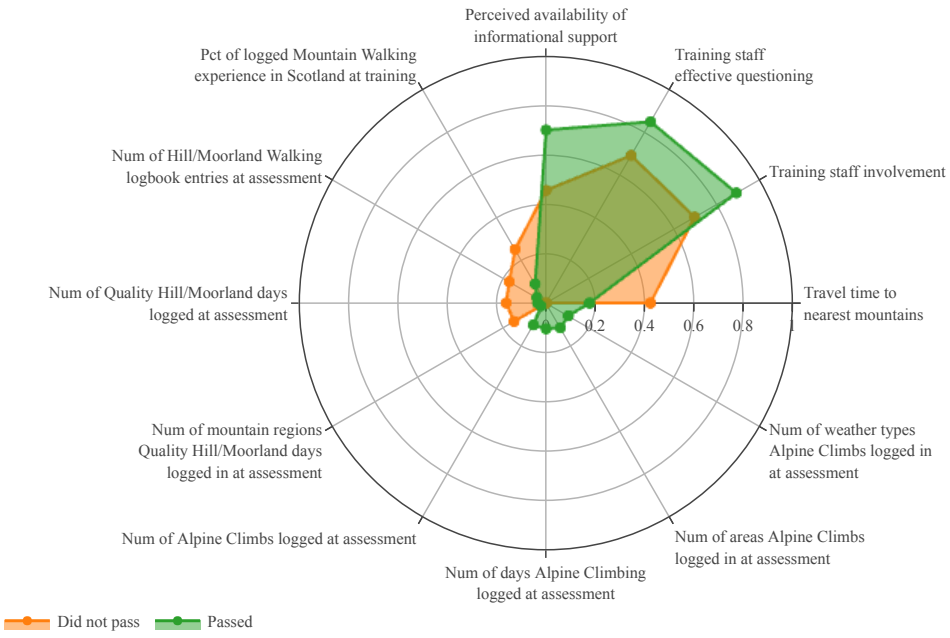

Figure 3.7: Centralised 2s RFE: Normalised training group means for candidates passing their first assessment - survey variables.

- feel that they have more informational support available to them
- log less experience below the standard for the Mountain Leader (Hill/Moorland walking) prior to assessment
- log more experience above the standard of the Mountain Leader (Alpine Climbing) prior to assessment

Table 3.11: Group 5 candidates passing first time, unstandardised group descriptive statistics

	Not passed			Passed		
Variable	mean	median	sd	mean	median	sd
Perceived travel time to the nearest mountainous region at registration[ab]	7.35	6	4.23	3.65	3	2.29
Perceived involvement of training staff[ab]	4.17	5	2.23	5.35	6	1.03
Total relationship conflict on the training course[a]	58.30	56	44.27	41.74	23	48.01
Perception of training course staffs' "observation" skills[a]	8.43	9	1.95	8.83	8	1.11
Perception of training course staffs' "effective questioning" skills[ab]	7.52	8	2.35	8.78	9	1.51
Perceived availability of informational support[ab]	2.83	3	1.15	3.80	4	1.44
Percentage of logged Mountain Walking experience in Scotland at training[b]	25.09	0	39.31	8.91	0	20.29
At assessment, number of...						
Hill/Moorland Walking logbook entries[b]	4.48	1	7.32	1.13	0	1.94
Quality Hill/Moorland days logged[b]	7.61	3	12.00	1.65	0	3.21
mountain regions Quality Hill/Moorland days logged in[b]	1.35	0	2.27	0.22	0	0.52
Alpine Climbs logged[b]	0.00	0	0.00	1.13	0	2.82
days Alpine Climbs logged on[b]	0.00	0	0.00	1.57	0	3.93
areas Alpine Climbs logged in[b]	0.00	0	0.00	0.35	0	0.88
weather types Alpine Climbs logged in[b]	0.00	0	0.00	0.52	0	1.27

[a] Included in Merged survey 2s 3s RFE

[b] Included in Centralised 2s RFE

3.5.3 Discussion

We retained two predictive models that could classify candidates as having passed their first assessment or not with good accuracy (up to 82.61%). None of the features selected as important discriminatory variables were included in the predictive models retained for the getting to assessment classification problems. There were 14 unique variables between the models, none of which were included in ether the female or male getting to assessment models. The features included in the final models can be considered under two broad headings, candidates experience of training and their preparation for assessment.

3.5.3.1 Experience of training.

Candidates who felt that course staff were more involved with their development, observed them more closely, and helped them to set goals and identify targets for attaining their goals were more likely to pass their first assessment than those who felt the course staff engaged in these behaviours to a lesser degree. Coaching usually aims to improve an individual's knowledge, skills, and competencies (Wagstaff et al., 2018). Course staff will engage in coaching behaviours to a greater or lesser extent and their facilitation of candidate's development will vary accordingly. Two coaching behaviours were selected as important discriminatory variables: effective questioning and goal setting. It is important to note that candidates who did not pass their first assessment reported that their course staff did engage in these behaviours, just not to the same extent as those who did pass.

Clear and specific goals are more effective than broad/vague goals for influencing behaviour change (Gould, 2005) and by closely observing a candidate's skills, course staff will have more information with which to help the candidate set goals. This close observation of candidates' skills would likely make them feel that their course staff were genuinely interested in their development. In addition, the use of effective questioning may encourage candidates to think and reflect on their actions, thus encouraging self-directed learning, thus supporting candidates' autonomy (Wagstaff et al., 2018).

We measured relationship conflict between candidates, and between candidates and staff on the training course and it was the total of these four items that discriminated between candidates who did and did not pass, suggesting that any conflict

experienced on a training course has a negative impact on the assessment outcome. Relationship conflict has generally been found to have negative relationships between both distal- and proximal-group outcomes (see de Wit et al., 2012). Whilst candidates on a Mountain Leader training course may only be a group for the five days of the course, they will share goals and be required to work together to complete tasks. It has been suggested that relationship conflict can distract groups from task accomplishment (Evan, 1965; Jehn, 1995). In the context of the Mountain Leader training course, this conflict may manifest as less time being spent on course content and the training staff coaching the candidates.

Candidates who felt they had more informational support available to them were more likely to have felt that they would have someone to give them advice about becoming a Mountain Leader and about performing at assessment if they needed it. In the sporting domain, perceived available support has been associated with positive effects on self-confidence and stress buffering (Rees and Freeman, 2007), performance (Freeman and Rees, 2009), and the processes underpinning performance (Rees et al., 1999). More specifically, perceived informational support has been correlated with greater situational control, increased challenge appraisal, decreased threat appraisal, and better performance (Freeman and Rees, 2009). When considering the findings of Study 1, it would be reasonable to expect that informational support may also help candidates to understand the, sometimes, fuzzy concept of "the standard" required for assessment. Understanding this would then help candidates better plan their preparation for assessment.

3.5.3.2 Preparation for assessment.

Interestingly, no QMD variables were selected in the final models, however, Hill/Moorland Walking, Mountain Walking, and Alpine Climbing variables were selected. Hill/Moorland Walking and Mountain Walking are types of mountaineering experience that do not meet the standard for a QMD because of the terrain that they cover, or because there has not been sufficient challenge to develop candidates Mountain Leader skills and therefore do not qualify as a QMDs. In contrast, Alpine Climbing is considered beyond the scope of the Mountain Leader qualification but does include terrain that will develop skills relevant to the Mountain Leader assessment (e.g., moving

safely in steep ground). One explanation for no QMD variables being included in the final models is that candidates with little experience may include all of their experience, whilst those who have a plethora of experience may only include enough in their logbook to meet the course prerequisite, leaving the majority of their experience unlogged. Thus, making it impossible to differentiate the highly experienced from those with little experience using QMD data alone; however, we would suggest that by using other DLOG variables it is possible to do just that.

The results presented above suggest that candidates who include Hill/Moorland Walking and Mountain Walking in their DLOG are less likely to pass their first assessment. Whist it is unlikely that this type of experience is detrimental to candidates' performance at assessment, its inclusion in their logbooks may indicate that these candidates have less experience and feel their logbook is weak, therefore they want to include as much experience as they can to "pad" their logbook out. In contrast, Alpine Climbing experience may be considered as an indicator of a candidate with more mountaineering experience—which may or may not be logged. These candidates may have logged this experience as it is important for higher level qualifications. These findings support the results of Study 1, which suggested that the quantity, quality, and variety of experience candidates had when they were assessed were all important for discriminating those who did pass from those who did not. Further, the findings suggests that having the right quality of experience is vital and that lower-quality experience is not a suitable substitute and that experience of mountaineering at a level higher than the Mountain Leader qualification may be beneficial to candidates.

Seven of the 23 candidates who did not pass their first assessment were only deferred because they had too few QMDs in their logbook at assessment.[7] It is important to highlight that the features presented here discriminate between candidates who do and do not pass their assessment, not between candidates who are and are not good enough to pass a Mountain Leader assessment, in terms of their skills and decision making. If we removed these particular candidates from the sample, we would have too few cases to perform the analysis, therefore, it is difficult at this juncture to answer the question "Is having more than the minimum experience beneficial for passing a Mountain Leader assessment?" If anything, the fact these candidates were able to

[7]Candidates who are deferred must be reassessed, however, when they are deferred for not having enough experience at assessment, their reassessment is simply presenting an assessor with an updated logbook.

perform at the standard required at assessment is evidence that one can pass the practical element a Mountain Leader assessment with fewer than 40 QMDs. Further, these candidate not passing only because of their logbooks is likely to be a contributory factor to the lower classification rates in the passing first time analyses than the getting to assessment analyses.

Candidates who live further from the mountains were less likely to pass; of the 12 candidates who lived more than 16 hours from the nearest mountains, only one passed their first assessment (8.33%), whereas of the 34 candidates who lived within six hours of the nearest mountains, 22 passed their first assessment (64.71%). Living further from the mountains may mean candidates are not able to travel to the mountains as frequently and may not be able to spend as long there when they do travel. This potential barrier may result in candidates logging more experience that is not as relevant (e.g., Hill/Moorland Walking) or not of the requisite quality (e.g., Mountain Walking) during their preparation for an assessment.

3.6 General Discussion

3.6.1 Overview

In this chapter, we aimed to identify important factors that discriminated candidates who (a) having been trained, went on to be assessed within 18 months of training from those who did not (for female and male candidates separately); and (b) having got to their first assessment, pass first time from those who do not. To achieve these aims we considered a wide range of potentially relevant variables (informed by the results of Study 1) and used pattern recognition analyses to analyse data collected from candidates. The use of the pattern recognition analyses allowed us account for the multifaceted nature of becoming a Mountain Leader by considering variables from several domains: personality, socio-demographics, intentions and expectations, motivation, experience of training, experiences post-training, social support, self-efficacy, and experience of assessment if relevant.

We were able to identify appropriate predictive models for all three classification problems. The results presented above suggest that there is no one single factor that is important for discriminating candidates, but there are combinations of factors that are

important. Whilst there is some overlap between the factors selected in the classification problems, there are factors unique to each problem. It is important to reiterate that the variables included in these models are the best discriminatory variables, therefore there may be other variables that are not included in the models but are very important for candidates getting to an assessment/passing their first assessment. For example, none of the variables included as important discriminatory variables are included in either of the first time pass models, but it is likely that in order to pass their first assessment, candidates must possess both the characteristics identified as important in the first time pass analyses *and* the characteristics identified as important for getting to assessment.

It is important that becoming a Mountain Leader fits into the wider context of a candidate's life in order for them to get to an assessment. This will allow them the time to prepare effectively for an assessment, likely resulting in them having higher levels of self-efficacy that they can become a Mountain Leader and perform tasks that will be required of them at assessment. Preparation for assessment seems particularly important for female candidates and training course staff may be able to help candidates maximise the benefits of their preparation by helping candidates to set personalised goals, which may be tailored to the candidate's life context. Further, it appears that it is important that this experience is varied and not all gained in the same area, but also that it is relevant to the Mountain Leader qualification. Experiencing changes in life may "derail" candidates' progress towards becoming a Mountain Leader and therefore could be considered as career turning points, especially if this change precipitates an alteration in the relative importance of becoming a Mountain Leader. It was important for candidates who did get to an assessment to have had a positive experience of their training course, that would help them to understand what they needed to do in order to prepare for an assessment, and then to have done so, (i.e., to have prepared effectively by gaining suitable and sufficient experience) in order to pass the assessment.

3.6.2 Potential Implications

If candidates leave their training course wanting to be assessed (i.e., they are motivated) and they are able to prepare effectively for an assessment they are more likely to be assessed. Any plan for getting to assessment that a candidate creates should consider how becoming a Mountain Leader fits into the rest of their life as this is clearly very

important. Such a process needs to include consideration of other life goals, how much time candidates feel they have available to prepare for an assessment, and how easy it is for them to access the mountains. It is important that these considerations are made, as if a plan is unrealistic it is unlikely to be followed and failure to attain goals will have a negative impact on candidates' motivation and confidence. Similarly, it is important that candidates who experience changes in their life reconsider their plan, as the changes they have experienced may then make the plan unrealistic.

Candidates' plans for consolidation should include goals that provide them with the opportunity for mastery experiences. When set in conjunction with course staff who can provide structure, these goals will offer candidates the opportunity to experience success as a result of good performance, which should increase their level of confidence to perform tasks related to passing a Mountain Leader assessment. Course staff may benefit from additional training, aimed at enhancing their coaching skills and helping them to provide a need supportive environment for their candidates.

3.6.3 Methodological Considerations

The strengths of this study include, the broad range of variables considered, the rigorous development of a survey tool to measure these variables, and advanced statistical analyses. However, we recognise that several limitations can be identified in this project. Firstly, most of the data used were collected retrospectively. Retrospective data will be less accurate as time increases between the event and when participants are sampled (Hopwood, 2013), and people may create their own narrative retrospectively which may or may not reflect reality. An example of this could be a candidate who did not pass their first assessment attributing their failure to the coaching (or lack thereof) they received on their training course. Given the complexity of some questions and the specificity of the time point that they related to, where possible, we used decomposition techniques like those found in The World Health Organization Health and Work Performance Questionnaire (HPQ; Kessler et al., 2003) to improve response accuracy and reduce the potential impact of recall bias.

Secondly, there is some evidence of sampling bias in the data used to identify the important discriminatory factors for both getting to assessment and passing. The proportion of female and male candidates who did get to assessment within 18 months

of their training course is not the same in the retrospective data (females = 23.21% and males = 41.35%) as it is in the population of candidates trained in the same period (females = 19.02% and males = 30.22%). In addition to this, the proportion of males who did not pass their first assessment is not the same in the retrospective data (13.50%) as it is in the prospective data (19.60%) or in the population[8] (19.80%); there is no evidence of the same problem in the data collected from female candidates. The simplest explanation for this is that candidates who are not assessed and male candidates who do not pass their first assessment are less likely to *retrospectively* respond to the survey. Whilst there may be a subset of candidates who are not represented in the data collected as part of this project, we believe that the findings presented in this chapter can be used to make a positive impact on the completion rate of the Mountain Leader qualification. This belief is based not only on the analyses of retrospective and prospective data presented here, but their congruence with the results from Study 1 and existing literature.

3.6.4 Future directions

The nature of the data and classification problems mean that as time passes, further analysis of the data is possible, as more candidates are further in time from their training course. For example, in the future, it will be possible to analyse the data to understand which factors are important for discriminating candidates who have been assessed five years post-training from those who have not been. We would also suggest that further analysis of these data in the future should go some way to mitigating the sampling bias mentioned above, as the response rate in the prospective data is similar to that in the population and reduce the impact of recall bias. However, a truly prospective study that collected data from candidates at registration, training, and during their consolidation would likely overcome the limitations described above.

The results presented in this chapter suggest that both experience and self-efficacy are important for both female and male candidates at various stages of the pathway. However, based on the number of relevant variables selected, both experience and self-efficacy appear to be more important variables for discriminating female candidates, rather than male candidates, who are assessed within 18 months of their

[8]Candidates who were first trained after 2016.

training course from those who are not. It would be useful for future studies to consider the nature of the relationship between experience and self-efficacy and potential gender-differences, this would not only be useful for developing the Mountain Leader qualification but may also be useful for better understanding the relative magnitude of the effects of positive and negative experiences on self-efficacy.

3.7 Conclusion

From up to 529 features, we were able to identify predictive models of no more than 16 variables, that could correctly classify candidates with up to 94.64% accuracy, based on three different outcome variables with good to excellent accuracy. The discriminatory variables included in these models covered several different temporal aspects of the training pathway and related to both candidates and course staff. This study supports the view of previous research that the development of expertise is multi-faceted and complex.

In order to become a Mountain Leader, these findings suggest that it is important that candidates feel that becoming a Mountain Leader is an important goal, are able to prepare effectively for an assessment including gaining relevant experience, feel confident that they can perform Mountain Leader related tasks at an assessment, and have a positive experience of the training course. These findings leave Mountain Training with several areas (e.g., training course staff to improve their coaching behaviours; helping candidates to set appropriate and realistic goals, thereby increasing their opportunities for mastery experiences and to make progress towards becoming a Mountian Leader) to focus their efforts if they wish to improve the completion rates of the Mountain Leader qualification.

Chapter 4

Self-Efficacy and Quality Mountain Days

4.1 Introduction

An individual's level of self-efficacy reflects their level of confidence that they can perform a specific task successfully at a given moment in time (Bandura, 1977). According to self-efficacy theory, if one possesses the necessary skills and is sufficiently motivated, their level of self-efficacy will be the primary determinant of performance, effort, and persistence—especially in the face of adversity (Bandura, 1977, 1982, 1997). Self-efficacy theory posits that there are four main sources of efficacy information, in order of decreasing influence on efficacy-beliefs: (1) previous performance accomplishments, (2) vicarious experience/modelling, (3) social/verbal persuasion, and (4) physiological/emotional states (Bandura, 1982). Evidence supports the influence of performance accomplishments on efficacy, with, a recent meta-analysis, Sitzmann and Yeo (2013) showing previous performance to have a moderate to strong effect on efficacy-beliefs ($\rho = .18 - .52.$). As self-efficacy theory has been introduced previously (see 1.4.3 General Introduction), we will not repeat details here.

The results of Chapter 2 suggested that it was important that both female and male candidates were confident in their skills in order for them to become Mountain Leaders, but that female candidates needed to be more confident than male candidates did before being assessed. Further, the results of Chapter 3 suggested that candidates' self-efficacy to perform specific skills relating to the Mountain Leader qualification were

important variables for discriminating both female and male candidates who did and did not get to an assessment within 18 months of their training course, although more self-efficacy variables were important for discriminating female candidates than were important for discriminating male candidates. In addition, the findings from Chapter 3 also highlight the importance of experience. Mountain Training require candidates to have a prerequisite level of specific experience before being assessed (40 Quality Mountain Days), and in Chapter 3, some experience related variables (e.g., QMDs 18 months post-training) were important for discriminating candidates who were assessed 18 months post-training from those who were not. As with the self-efficacy variables, more experience related variables were important for discriminating female candidates who were assessed from those who were not than they were for male candidates.

Whilst there are no previous studies on gender differences in self-efficacy within the Mountain Leader community, research in education has examined the effects of gender on self-efficacy for over 30 years and has reported that females have lower levels of self-efficacy than males (e.g., Murphy et al., 1989; Klassen and Chiu, 2010). The field of computer education in particular has considered the relationship between experience, gender, and self-efficacy. For example, several studies have examined the effects of gender and training on self-efficacy (e.g., Cassidy and Eachus (2002), Murphy et al. (1989), Torkzadeh and Koufteros (1994)). Torkzadeh and Koufteros (1994) considered the effects fo gender and training on four factors of computer self-efficacy: beginning skills, mainframe skills, advanced skills, and file management. Torkzadeh and Koufteros (1994) found evidence for gender differences pre-training for one factor of self-efficacy; females reported lower self-efficacy scores on the file management factor pre-training. Both female and male participants reported increased levels of self-efficacy post-training for all four factors and there was no longer a significant difference in their scores for any factor.

It is interesting that the difference in file management self-efficacy between female and male participants was no longer significant post-training, suggesting that there may be an interaction between gender and training time point on file management self-efficacy, where the effect of training was more positive for female participants (although the authors did not test for the presence of interactions directly). In addition, results reported by Cassidy and Eachus (2002) suggest that male participants have

higher level of self-efficacy pre- and post-training, but do not appear to control for previous experience in their analyses, despite reporting that male candidates "were more experienced."However, when examining these articles, it would appear that an important variable—experience—has potentially been overlooked when considering gender differences in self-efficacy.

When considering the results of the previous two chapters through the lens of self-efficacy theory, it is apparent that the relationship between confidence and experience is a likely to influence the likelihood of a candidate becoming a Mountain Leader. Further, it is likely that this relationship is more important for female than male candidates. Based on the results of the previous chapters and the computer education studies presented above, it seems reasonable to hypothesise that (a) candidate experience will positively predict their level of self-efficacy because experience increases the opportunity for performance accomplishments, and (b) that this relationship will be influenced by gender. More specifically, this positive relationship should be stronger for female candidates than males.

Therefore, with this interaction perspective in mind, the aim of the present chapter was to investigate the relationship between experience, and gender differences on Mountain Leader related self-efficacy. To do so, we used the data collected for the work reported in Chapter 3 and Appendix B. The remainder of this chapter is structured as follows. Firstly, we present Study 5, which reports on the development and initial validation of a measure of self-efficacy for skills related to becoming a Mountain Leader using participants from Appendix B. Following that, in Study 6, we used the measure created in Study 5 to examine the relationship between self-efficacy, experience, and gender using participants from Chapter 3. Participants in Appendix B completed their first Mountain Leader training course between 2008 and 2016, whereas participants in Chapter 3 completed their first Mountain Leader training course in either 2017 or 2018. The availability of digital logbook (DLOG) data is greater the more recently a candidate was trained. Therefore, we used the data from participants in Appendix B for Study 5 and data from participants in Chapter 3 for Study 6. Finally, we discuss these two studies and consider future directions for this research.

4.2 Study 5: The Mountain Leader Self-Efficacy Scale.

Self-efficacy is domain specific. Therefore, it is important that any measure of self-efficacy is also domain specific (Bandura, 2006). There is no existing measure of Mountain Leader self-efficacy, however, in Chapter 3 we developed 11 self-efficacy items for skills that candidates would be required to perform at assessment based on the Mountain Leader candidate handbook and syllabus (Mountain Training UK, 2015a), and a separate skills checklist (Mountain Training UK, 2015b). Therefore, the aim of Study 5 was to develop self-efficacy measure using these 11 items and provide inital evidence for its validity.

4.2.1 Methods

4.2.1.1 Participants.

Participants for Study 5 were 526 candidates who had attended a Mountain Leader training course between 2008 and 2016 and completed the survey used for the Chapter 3 pilot work (21.67% female, M_{age} 39.14 years, $SD = 11.98$, 72.24% had been assessed when completing the survey).

4.2.1.2 Measures.

The initial Mountain Leader Self-efficacy Scale (MLSS) comprised 11 items developed by the first author using a inductive content analysis (Cho and Lee, 2014) of the candidate handbook and syllabus (Mountain Training UK, 2015a) and skills checklist (Mountain Training UK, 2015b). These items were then checked with the second and third authors, with any disagreements being discussed until a consensus was reached. Finally, Mountain Training viewed the items and agreed that they provided good coverage of the skills that would be covered on an assessment and were worded in a way which would be understood by their candidates. The final scale was made up of 11 items (e.g., "lead a group effectively in the mountains") rated on a scale of *could not do at all* (0) to *highly certain could do* (100) with a mid-point anchor (*moderately could do*; 50). For those candidates who had been assessed, we asked them to think about how

confident they were when they arrived for assessment. We asked candidates who had not been assessed how confident they were when completing the survey.

4.2.1.3 Analyses.

We used an exploratory approach to confirmatory factor analysis with a robust maximum-likelihood estimator using the `R` package `lavaan` (Rosseel et al., 2020). A model was considered a good fit to the data if the Yuan-Bentler (Y-B; Yuan and Bentler, 1997) χ^2 test was not-significant. However, given that this is a test of exact fit, many researchers suggest that it is overly conservative and is often significant when performing CFAs on large samples. Thus, Jöreskog (1989) suggests inspecting the χ^2/df ratio to assess model fit, with values < 2 indicating good model fit. In addition to the χ^2/df ratio, we also assessed several approximate fit indices and considered the model an approximately good fit if it satisfied the following criteria: the comparative fit index (CFI; Bentler, 1990) was greater than or equal to .95, the root-mean-square-residual (RMSEA; Steiger, 1990) was less than or equal to .06, and the standardised root-mean-square-residual (SRMR) was less than or equal to .08 (Hu and Bentler, 1999). We assessed internal consistency using composite reliability (ω), values >.70 are considered acceptable (Fornell and Larcker, 1981).

In addition to examining the factor structure, we sought to test the factor solution for measurement invariance between female and male candidates. Testing for measure invariance allows one to examine whether scores of a construct have the same meaning in different conditions (Meade and Lautenschlager, 2004). In this chapter, we tested if the MLSS scores had the same meaning for female and male candidates. More specifically, we sought to test for *configural invariance* (is the factor structure the same across groups?), *weak invariance* (are the factor loadings the same across groups?), and *strong invariance* (are the item intercepts the same across groups?). To do so, we specified three models with additional constraints in each model, which increases the model degrees of freedom.

To test for configural invariance, the same CFA model is tested in each group, if this model is inconsistent with the data then the measure is considered non-invariant at all levels. The hypothesis of weak invariance in this chapter is that the MLSS unstandardised factor loadings are the same for female and male candidates and the

hypothesis of strong invariance is that the MLSS unstandardised factor loadings and item intercepts are the same for female and male candidates (cf. Kline, 2016). These three models are considered to be *nested models* as they have the same model structure but test increasingly strict hypotheses (Kline, 2016). Therefore, changes in fit can be attributed to the constraints associated with each level of invariance. Current guidelines suggest that researchers should use multiple fit statistics to assess model fit (Kline, 2016) and that χ^2 and at least two other fit indices should be reported when assessing the fit of invariance models (Putnick and Bornstein, 2016). In this study we used CFI, RMSEA, and SRMR. Simulation results from Chen (2007) suggest that $\Delta_{CFI} > .005$ and $\Delta_{RMSEA} > .010$ between more and less constrained models are adequate for detecting non-invariance. Therefore, when testing each measurement invariance hypothesis, we considered models with increases in CFI and RMSEA greater than these values to be non-invariant at the relevant level of invariance.

4.2.2 Results

Results of a single factor CFA revealed a poor fit, (Y-Bχ^2 = (44) 310.87, p < 0.01, CFI = 0.82, RMSEA [90% CI] = 0.18 [0.20,0.17], SRMR = 0.07). We subsequently made modifications to the model in an iterative fashion based on examination of standardised loadings, modification indices and theoretical considerations (cf. Biddle et al., 2001). The first modification made to the model was creating two separate factors. We created two factors as the modification indices suggested that there was covariance between two of the items which was not adequately explained by the model. These two items were about *emergency skills* (i.e., those only required in the event of an emergency), whilst the remainder of the items reflected *routine skills* (i.e., those required on a routine basis)). Given the difference in theoretical focus across these two factors we deemed this distinction appropriate, and hence created a new two-item factor with the aim of better modelling this covariance.[1]

The remaining modifications involved the removal of items based on their modification indices (n = 5), which suggested that there was covariance unaccounted for in the model, but there was no clear theoretical justification for creating any further

[1]Whist a single factor CFA model must have at least three indicators for the model to be identified, when the number of factors is greater than one, each factor must have at least two indicators when there are no correlated error terms (Kline, 2016).

factors, hence the items were removed. This process led to the retention of a six-item, two-factor model, which was a good fit to the data and displayed good internal consistency (see Table 4.1). To test the discriminant validity of the two factors, we performed a Satorra and Bentler (2001) scaled difference χ^2 test (S-Bχ^2_{diff}) on the two-factor model and a respecified one factor model. The results of this test supported the discriminant validity of the two-factor model (S-Bχ^2_{diff} (1) = 356.42, p < .001)).

Because we wished to create a measure that could be used to assess self-efficacy of female and male candidates, in order to make any comparisons between these two groups it was important that at least the hypotheses of configural and weak invariance were satisfied. The results of the measurement invariance analyses carried out on the six-item two-factor model supported the hypotheses of configural, weak invariance, and strong invariance (see Table 4.2). These findings suggest that the scores for the two MLSS factors have the same meaning for female and male candidates.

Table 4.1: Factor loadings and model fit indices for the two-factor MLSS in Study 5 and 6.

	Study 5	Study 6
Routine skills		
Wild camp for two nights in any weather.	.62	.62
Look after myself and others in steep ground/crossing a river.	.80	.77
Navigate to a chosen point on a map in any conditions, night or day.	.79	.81
Plan a mountain day that is appropriate for the group.	.80	.83
Emergency skills		
Provide immediate medical care in the mountains.	.91	.90
Respond appropriately to an emergency (e.g., broken leg).	.96	.98
Composite reliability (ω)		
Routine	.84	.85
Emergency	.93	.94
Inter-factor correlation		
Routine-Emergency	.71	.70
Two-factor model fit indices		
Y-B χ^2	7.35	7.35
df	8	8
p	0.50	0.13
CFI	1.00	1.00
RMSEA [90% CI]	0.00 [0.00,0.08]	0.05 [0.00,0.10]
SRMR	0.02	0.02

Note:
Y-B = Yuan-Bentler. CFI = comparative fit index. RMSEA = root-mean-square error of approximation. SRMR = standardised root-mean-square residual.

Table 4.2: MLSS measurement invariance results for Study 5 and 6.

Model	Y-Bχ^2	df	p	ΔX^2	Δdf	p	CFI	ΔCFI	RMSEA [90% CI]	ΔRMSEA	SRMR	ΔSRMR
Study 5												
Configural	21.05	16	.18	NA	-	NA	.994	NA	.050 [.000,.103]	NA	.025	NA
Weak	26.01	20	.17	4.93	4	.29	.993	-.001	.049 [.000,.096]	-.002	.037	.012
Strong	30.30	24	.18	3.56	4	.47	.994	.000	.044 [.000,.086]	-.005	.038	.001
Study 6												
Configural	24.08	16	.09	NA	-	NA	.991	NA	.062 [NA,.109]	NA	.030	NA
Weak	26.80	20	.14	2.86	4	.58	.993	.001	.051 [.000,.096]	-.011	.036	.007
Strong	36.71	24	.05	11.96	4	.02	.987	-.006	.062 [.008,.100]	.011	.042	.005

Note:
All fit indices estimated using robust SEs.
The $\Delta\chi^2$ is a robust difference test that is a function of two standard (not robust) statistics.

4.2.3 Discussion

The aim of Study 5 was to develop a measure of self-efficacy for skills related to becoming a Mountain Leader—the MLSS—using data that were previously collected as part of Chapter 3. Fit indices and standardised factor loadings suggested that the six-item two-factor model fit the data well. In addition, the hypotheses of configural, weak, and strong invariance were supported, which suggests that the MLSS factor factors are not different for female and male candidates. Having established a factor structure of the MLSS using an exploratory approach to CFA, in Study 6 we sought to test the proposed factor structure using a strictly confirmatory approach to CFA.

4.3 Study 6

The aim of Study 6 was twofold: (a) to confirm the factor structure of the MLSS and (b) to test the hypotheses presented in the introduction to this chapter.

4.3.1 Methods

4.3.1.1 Participants.

Participants for Study 6 were a new sample of 431 candidates who had attended a Mountain Leader training course in 2017 or 2018 and completed the survey used for Chapter 3 (35.27% female, M_{age} 39.14 years, $SD = 11.93$, 29.93% had been assessed when completing the survey).

4.3.1.2 Measures.

We used the six-item MLSS from Study 5 to measure *routine skills self-efficacy* and *emergency skills self-efficacy*. In addition, we operationalised experience as either the sum of the number of QMDs that a candidate had when answering the survey, for candidates who had not been assessed when completing the survey, or the number of QMDs at assessment for those who had been assessed. We chose the number of QMDs as the measure of experience as that is the type of experience Mountain Training requires candidates to accrue in order to become a Mountain Leader.

4.3.1.3 Analyses.

To test the fit and invariance of the two-factor model retained from Study 5, we carried out a CFA and measurement invariance analyses as specified for Study 5. Following this, we used moderated hierarchical regression analyses to test the interactive effects of gender and experience on self-efficacy. We examined each factor of self-efficacy separately. To obtain scores for the two self-efficacy variables, we retained the factor scores from the two-factor CFA; factor scores are a better estimate of the true value of the latent construct than a sum-scores, as factor scores account for measurement error (cf. Grice, 2001).

For each factor, using the factor scores as the dependent variable, we fitted three regression models to the data. The first model (Step 1) had gender as the sole predictor of the dependent variable. In the second model (Step 2), we included the main effects of both gender and experience as predictors of the dependent variable. Finally, for the third model (Step 3), we included both the main effects and interactive effect of gender and experience as predictor variables. Alpha was set at .05 and we standardised all continuous variables before entering them into the regression models to provide a common metric, thus aiding the interpretation of the interaction term (Aiken and West, 1991).

4.3.2 Results

4.3.2.1 MLSS Model Fit and Invariance.

The two-factor model was an approximately good fit to the data (see Table 4.1) and the results of a S-Bχ^2_{diff} test supported the discriminant validity of the two-factor model when compared to a single factor model (S-Bχ^2_{diff} (1) = 355.38, p < .001). The results of the measurement invariance analysis provided good support for the hypotheses of configural and weak invariance, however, there was evidence to reject the hypothesis of strong invariance. The $\Delta\chi^2$, ΔCFI, and $\Delta RMSEA$ were all just outside the criteria specified above for detecting invariance. Rejecting the hypothesis of strong invariance suggests that there is a *differential additive response style* (i.e., female and male candidates do not use the response scale in the same way) and that differences in group means should be interpreted with caution as there may be a systematic cause for the

difference in scores that is not explained by the MLSS factors (Kline, 2016). Whilst this may influence gender differences in self-efficacy, this should not change our interpretation of the main effect of experience on self-efficacy, nor the interactive effect of gender and experience on self-efficacy.

4.3.2.2 Self-efficacy, Experience, and Gender.

Before interpreting the results of the regression analyses, we used the variance inflation factor (VIF) to assess the multicollinearity of the predictor variables. Values close to 1 are preferred and values less than 5 are acceptable (Tabachnick and Fidell, 2007); the maximum VIF in these data was 1.35. Descriptive statistics for study variables are presented in Table 4.3.

Table 4.3: Descriptive statistics and correlations between study variables (N = 433.)

Variable	M	SD	1	2	3
1. Gender[a]	0.35	0.48			
2. Experience	34.06	25.97	-.12*		
3. Routine self-efficacy	0.33	8.92	-.31**	.27**	
4. Emergency self-efficacy	0.70	18.37	-.14**	.20**	.75**

Note:

* indicated p < .05. ** indicates p < .01.

[a] Males coded as 0 and females coded as 1.

4.3.2.2.1 Routine Skills. Gender predicted self-efficacy at Step 1 ($b = -5.75$, 95% CI [−7.43, −4.07], $t(431) = -6.72$, $p < .001$), with female candidates having lower levels of self-efficacy to perform routine skills than males. Experience predicted self-efficacy at Step 2 over and above gender ($b = 2.11$, 95% CI [1.33, 2.90], $t(430) = 5.30$, $p < .001$, $\Delta R^2 = .06$), with greater experience being associated with greater levels of efficacy. Of more interest, the interaction was significant ($b = 2.31$, 95% CI [0.49, 4.12], $t(429) = 2.50$, $p = .013$, $\Delta R^2 = .01$). In line with our hypotheses, simple slope analyses suggested that the positive relationship between experience and self-efficacy was stronger for female candidates (b = 3.86, p 0.01) than for male

candidate (b = 1.55, p < 0.01). Figure 4.1 shows the nature of the interaction.

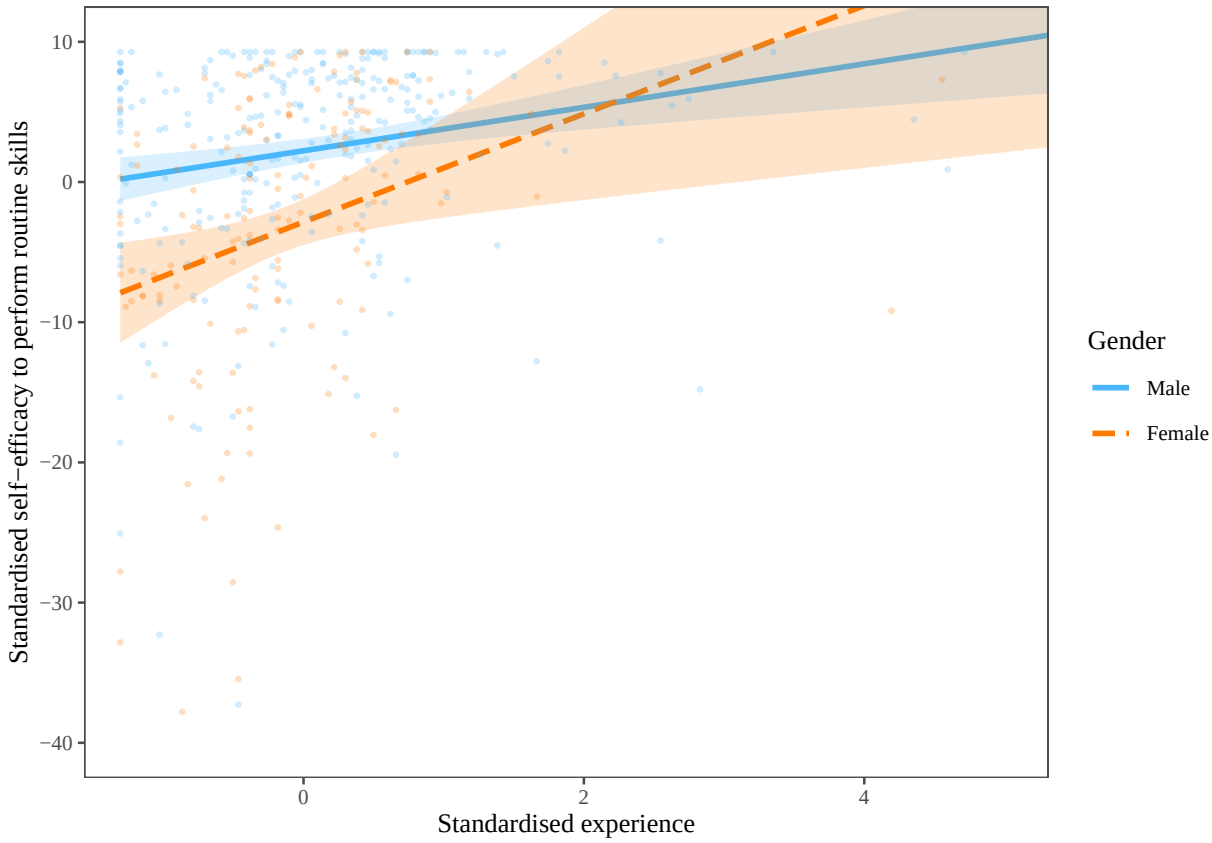

Figure 4.1: Interactive effects of gender and experience on self-efficacy to perform routine Mountain Leader skills N = 433. Ribbons represent the 95% CI.

4.3.2.2.2 Emergency Skills. Gender predicted self-efficacy at Step 1 ($b = -5.23$, 95% CI [−8.84, −1.63], $t(431) = -2.85$, $p = .005$), with female candidates again having lower levels of self-efficacy to perform emergency skills than males. As with the previous analysis, experience predicted self-efficacy at Step 2 over and above gender ($b = 3.54$, 95% CI [1.84, 5.24], $t(430) = 4.09$, $p < .001$, $\Delta R^2 = .04$). However, the interaction term was not significant in Step 3 ($b = 2.99$, 95% CI [−0.97, 6.95], $t(429) = 1.48$, $p = .139$, $\Delta R^2 = .00$).

Table 4.4: Regression analyses examining interactions between gender and experience on self-efficacy to perform routine and emergency Mountain Leader skills.

Predictor	b	b 95% CI	sr2	sr2 95% CI	Fit	Difference
Routine						
Step 1						
(Intercept)	2.35**	[1.35, 3.34]			R^2 = .095**	
Gender:Female	-5.75**	[-7.43, -4.07]	.09	[.05, .15]	95% CI[.05,.15]	
Step 2						
(Intercept)	2.18**	[1.22, 3.15]				
Gender:Female	-5.29**	[-6.93, -3.65]	.08	[.03, .13]	R^2 = .150**	Δ R^2 = .055**
Experience	2.11**	[1.33, 2.90]	.06	[.02, .10]	95% CI[.09,.21]	95% CI[.02, .10]
Step 3						
(Intercept)	2.23**	[1.26, 3.19]				
Gender:Female	-5.09**	[-6.72, -3.45]	.07	[.03, .12]		
Experience	1.55**	[0.65, 2.45]	.02	[-.00, .05]	R^2 = .163**	Δ R^2 = .012*
Gender:Female X Experience	2.31*	[0.49, 4.12]	.01	[-.01, .03]	95% CI[.10,.22]	95% CI[-.01, .03]
Emergency						
Step 1						
(Intercept)	2.53*	[0.40, 4.67]			R^2 = .019**	
Gender:Female	-5.23**	[-8.84, -1.63]	.02	[.00, .05]	95% CI[.00,.05]	
Step 2						
(Intercept)	2.26*	[0.16, 4.36]				
Gender:Female	-4.46*	[-8.02, -0.90]	.01	[-.01, .03]	R^2 = .055**	Δ R^2 = .037**
Experience	3.54**	[1.84, 5.24]	.04	[.00, .07]	95% CI[.02,.10]	95% CI[.00, .07]
Step 3						
(Intercept)	2.32*	[0.22, 4.42]				
Gender:Female	-4.19*	[-7.77, -0.62]	.01	[-.01, .03]		
Experience	2.82**	[0.86, 4.77]	.02	[-.01, .04]	R^2 = .060**	Δ R^2 = .005
Gender:Female X Experience	2.99	[-0.97, 6.95]	.00	[-.01, .02]	95% CI[.02,.10]	95% CI[-.01, .02]

Note:
A significant b-weight indicates the semi-partial correlation is also significant.
b represents unstandardised regression weights.
sr2 represents the semi-partial correlation squared.
Square brackets are used to enclose the lower and upper limits of a confidence interval.
* indicates $p < .05$. ** indicates $p < .01$.

4.3.3 Discussion

Study 6 had two aims, firstly, to confirm the factor structure of the MLSS and secondly, to test the two hypotheses presented in the introduction to this chapter. Study 6 confirmed the factor structure of the MLSS and its weak invariance for female and male candidates. In addition, findings provided evidence that: female candidates are less confident than their male counterparts; increased levels of experience predict increased levels of self-efficacy, supporting hypothesis one; and that the relationship between experience and self-efficacy to perform routine skills is more positive for female candidates than it is for male candidates, providing partial support for hypothesis two.

Experience may have a greater positive effect on efficacy to perform routine skills for female candidates than male candidates due to female candidates with less experience having lower levels of self-efficacy than their male counterparts. Indeed, the results of Study 1 suggest that, in some instances, female candidates felt that in order to feel confident enough to be assessed they needed to do more to prepare for an assessment (see Section 2.3.1.1.1.1). When considering the results of Study 6, specifically female candidates with little experience being less confident in their routine skills than male candidates of equal experience, one may interpret the findings of Study 1 differently. That is rather than female and male candidates having different thresholds of confidence that they must meet before being assessed, they in fact have different initial levels of confidence and require experience to gain enough confidence to feel ready to be assessed.

The most parsimonious explanation for why there was no interactive effect evident for emergency skills is because of the few opportunities Mountain Leader candidates get to practice their emergency skills when gaining QMDs, at least in comparison to the opportunities to practice routine skills. Accidents in the mountains are relatively rare; between 2009 and 2016 it is estimated that ~4,000,000 people visited Snowdon[2] yet only 1,081 people required assistance from mountain rescue (i.e., $< 0.01\%$ of visitors; Snowdonia National Park Authority, 2017). It is possible that the scarcity of experience in emergency situations means that any experience relating to emergencies is equally valuable for female and male candidates. Future studies of the effects of experience on self-efficacy may consider the scarcity of specific types of experience on their relative contributions to self-efficacy.

[2]Snowdon is the highest mountain in England and Wales and one of the busiest in the world.

Higher levels of experience predicted higher levels of self-efficacy, however the proportion of variance in the MLSS factors explained by experience (i.e., the number of QMDs accrued) was small. There are several explanations for the modest size of this relationship. One such explanation is that whilst QMDs are important, they are not the only source of self-efficacy. This suggestion is concordant with self-efficacy theory, which suggests previous experience is only one source of efficacy beliefs. In addition, it is likely that other forms of experience are important in establishing efficacy beliefs (e.g., scrambling experience). Another factor that is likely to influence the variance explained is the variability in accuracy of candidates' logbooks. As noted in Chapter 3, not all candidates will use their logbooks in the same way, thus adding noise to the data. In addition to the different use of logbooks, although the number of QMDs is an important real-world measure of experience, it does remain a somewhat crude measure of experience. More specifically, simply reporting the number of QMDs does not consider other factors that would, according to self-efficacy theory, be important in shaping efficacy-beliefs (e.g., was the experience perceived as a success or a failure? What were the candidates' attributions of the experience? Was the experience challenging or easy?).

Given that the number of QMDs in a candidates logbook is quite a crude measure of there experience for the reasons listed above, one may consider the recovery of any effect from the noise resulting from the measures crudeness to be surprising. Therefore, the fact that additional variance was explained by the interaction between gender and experience on routine self-efficacy is noteworthy. It is also important to consider that whilst the number of QMDs logged may be a crude measure of candidates' experience, it is the measure that Mountain Training use and it is therefore useful for them to understand the efficacy of that measure. To overcome the limitations identified in this study, a future prospective longitudinal study that measures self-efficacy using the MLSS and collects more accurate data about candidates' experiences would likely provide a better understanding of the relationships examined in Study 6. We hypothesise that in such a study, experiences that were: perceived as successful, attributed internally (e.g., having tried), and appropriately challenging would have the greatest positive influence on self-efficacy beliefs. In contrast, experiences that were perceived as a failure would decrease self-efficacy beliefs, especially when coupled with an internal attribution.

At this juncture we should note that these results should be interpreted with

some caution as in Study 6 the hypothesis of strong invariance was rejected. We recognise that the changes in fit indices when testing the hypothesis of strong invariance are (marginally) greater than those specified above for detecting non-invariance. However, given that several researchers caution against the use of exact cut-off values or "golden rules" (e.g., Kline, 2016; Markland, 2007) and one hypotheses of this study was that there would be an external variable—experience—that had a different effect on self-efficacy for female and male candidates we suggest that this violation of the hypothesis of strong invariance may be due to a difference in the mean experience of female and male candidates ($\Delta M = 0.22$, 95% CI [0.04, 0.40], $t(380.44) = 2.35$, $p = .019$). In addition, strict measurement invariance (as examined in this study), could be considered overly restrictive (cf. Muthén and Asparouhov, 2012) and an *approximate measurement invariance* approach could be tested in the future to better understand the exact nature of the non-invariance of item intercepts. Approximate measurement invariance is carried out using Bayesian structural equation modelling and involves specifying small (i.e., *approximately* zero) prior variances for parameters that would be fixed to zero in normal structural equation models (including CFAs). This change therefore allows models to better deal with unimportant levels of model miss-fit (cf. Muthén and Asparouhov, 2012; van de Schoot et al., 2013a).

Another explanation for the hypothesis of strong invariance being met in Study 5 but not in Study 6 could be the difference in the proportion of candidates in each study that had been assessed when answering the surveys (72.24% and 29.93% respectively). In addition to the prerequisite of having 40 QMDs logged prior to assessment, candidates must also hold a valid first aid certificate. Whist not having data to support this hypothesis, we would suggest that more participants in Study 5 had received first aid training than in Study 6 and that first aid training is likely to influence an individual's efficacy to perform the skills in the emergency skills factor (i.e., provide immediate medical care in the mountains and respond appropriately to an emergency [e.g., broken leg]).

4.4 Applied Implications

There are several implications of this work that are relevant to Mountain Training and Mountain Leader candidates. The MLSS comprises two distinct factors and items from both factors were selected as important discriminatory variables in Chapter 3. Therefore, it is likely to be important that candidates consider both sets of skills when preparing for an assessment. The results of Study 6 suggest that experience predicts self-efficacy for both sets of skills, therefore candidates should gain experience whilst preparing for an assessment. However, as a measure of experience, QMDs explain a relatively modest proportion of the variance in self-efficacy scores. It is likely that this in part due to inaccuracies in DLOGs, but it is also likely that other forms of experience are important influences on candidates' self-efficacy. Therefore, Mountain Training may wish to expand the prerequisites for experience to include sources of efficacy that are likely to be important (e.g., a measure of experience in steep ground).

Finally, when considered together, the results of Studies 3, 4, and 6 suggest that it is particularly important that female candidates seek opportunities for mastery experiences should they wish to become Mountain Leaders. Study 3 showed that it is important that female candidates are confident in a number of skills before being assessed. Study 6 showed that experience has a stronger relationship with self-efficacy for female candidates than it does for male candidates. and the findings of Study 4 indicate that it is important that candidates have sufficient relevant experience in order to pass their first assessment. Training course staff are in an ideal position to help candidates establish goals as they should understand candidates' abilities and which activities would provide appropriate challenge for given a candidate's abilities. We suggest that course staff helping candidates set clear and specific goals that are appropriate for them would help build their self-efficacy. In addition, supporting candidates to meet these goals during the consolidation phase would also be beneficial to candidates. Given the importance of these goals, it is likely that providing training to course staff to improve their goal-setting skills would be beneficial.

4.5 Future Directions

To our knowledge, previous research has not considered the interactive effects of experience and gender on self-efficacy, but in some areas has suggested that differences in self-efficacy is a result of "perceived masculinity" (Cassidy and Eachus, 2002, p 135). The equivocal results in Study 6 of the interactive effects of experience and gender on self-efficacy suggest that there may be other important factors to consider (e.g., differential availability/scarcity of types of experience). Given these results and the "real world" cross-sectional nature of the data in this chapter, future research should further investigate the interactive effects of experience and gender on self-efficacy both in a prospective longitudinal fashion and in an experimental fashion. Such studies should also consider the other variables identified as relevant (e.g., attributions) and may also consider relevant personality variables (e.g., emotional stability).

The MLSS was a good fit to the data in both Study 5 and Study 6 and can be used as a relatively short measure of two types of self-efficacy related to becoming a Mountain Leader. However, there are some items that could be improved. For example, one routine skill item reads, "look after myself and others in steep ground/crossing a river." This item could be split into four separate items: one about candidates looking after themselves in steep ground, one about candidates looking after themselves when crossing a river, one about candidates looking after others in steep ground, and one about candidates looking after others when crossing a river. Whilst increasing the number of items, improving the items may create a better measure of self-efficacy related to becoming a Mountain Leader, which could prove useful if Mountain Training wanted to better understand the relationship between experience and self-efficacy to perform specific skills. In addition, carrying out approximate measurement invariance studies may shed light on differences in use of the MLSS response scale by female and male candidates.

4.6 Summary and Concluding Discussion

In this chapter, we sought to create a measure of self-efficacy for skills related to becoming a Mountain Leader—the MLSS—and to examine the additive and interactive effects of experience and gender on self-efficacy as measured by the MLSS. The MLSS

developed in Study 5 and tested in Study 6 provided an appropriate fit to the data in both studies and the hypothesis of weak invariance between female and male candidates was supported, whilst the stricter hypothesis of strong invariance was only supported in Study 5. This issue notwithstanding, the model fit was good in both studies—which had different proportions of candidates who had been assessed—providing initial evidence for the validity of the measure for candidates who have and have not been assessed. The measurement invariance findings suggest that whilst the two self-efficacy factors have the same structure in both female and male candidates, their scores on the items are non-invariant. The moderated regression analyses in Study 6 provide evidence that: female candidates are less confident than their male counterparts, increased levels of experience predict increased levels of self-efficacy, and that the relationship between experience and self-efficacy to perform routine skills is more positive for female candidates than it is for male candidates.

In summary, candidates with more experience are more confident in their skills to perform tasks related to becoming a Mountain Leader and it is particularly important that female candidates gain relevant experience in order to be confident in their routine skills.

Chapter 5

Impact Activities

- MTUKI
  - October 2018 workshop
  - October 2019 workshop
  - Coaching workshop (November?) 2019
- MTE October 2018 workshop
- MTC
  - October 2019
- MTS
  - October 2019
  - Provider workshop
  - Reanalysis of PhD data (see Appendix D)

Chapter 6

General Discussion

We have finished a nice book.

Appendix A

Mountain Training Executive Officer Interview Guide

A.1 Interview A

Thank you for agreeing to be interviewed today. The information that I sent you earlier outlined five themes to cover over the two interviews; during this interview, I would like to cover four of them: Candidate Background, Candidate Career History, Social Influence, and Personal Characteristics.

As the purpose of these interviews is to understand which factors differentiate those who do become Mountain Leaders from those who do not, I will ask you about "completion". Someone who "completes" is someone who has passed a Mountain Leader assessment. However, I am interested in more than just who passes their assessment and who does not. Those who do not complete, may have attended been assessed but not passed; been trained or not assessed; or even registered and not trained.

At this early stage of the project it is important that we cover as many different variables as possible. As such, some of the things I ask you about may not seem immediately relevant to completion of the Mountain Leader award, or may seem very similar to other questions that I have asked. It would be much appreciated if you would answer each question as well as you can.

This interview should last between 2 and 3 hours, there is a break scheduled into it. As stated on the information sheets I will be recording the interview, I will also write some things down during the interview for me to refer back to, please don't let this

distract you.

One of the challenges of working with a small group of participants is ensuring that their confidentiality is preserved. Holding a public position can make maintaining confidentiality more challenging. I have considered this issue and various ways of overcoming this challenge at great length and intend to work closely with you to reach the highest levels of confidentiality possible. You have already read and signed the confidentiality and consent agreements prior to this meeting, which outline the process for maintaining confidentiality. I would however like to reiterate that you will be given opportunities, throughout the process, to review the information you have provided and raise any concerns you may have.

Stop me at any point if you would like a break and please ask for clarification if anything is not clear.

A.1.1 Candidate Background.

I'm going to start by asking you some questions about the Mountain Leader candidates who attend courses in your region to develop an idea of who they are and where they come from. I make no assumptions about what is "good" or "bad" with regards to becoming a Mountain Leader, so there are certainly no "right" or "wrong" answers. Whatever you say in this interview is between you and the research team, I would like you to speak freely about the Mountain Leader award and its candidates in your region.

- Could you start by describing a typical group of six candidates on a Mountain Leader training course to me?
- Can you give me some examples of the types of people who are always candidates who you will always see on a Mountain Leader award course in your region?
  - Prompt: Age, sex, location, ethnicity, socio-economic
- Can you give me some examples of other candidates who you see on Mountain Leader award courses in your region?
  - Prompt: Age, sex, location, ethnicity, socio-economic
  - …and do these candidates go on to become mountain leaders?
- Are there any types of people who are notably absent from Mountain Leader award courses in your region?

  - Prompt: Age, sex, location, ethnicity socio-economic

- Does the make up of an assessment course looks the same? If not, how does it vary?
  - Are there any differences in the numbers of men and women on training and assessment courses? If so, why do you think that is?
  - Are completion rates for men and women the same? If not, why do you think that is?

- How, if at all, do you think that a candidate's age at registration affects their progression through the Mountain Leader award?
  - What is it about being older or younger that makes a difference?
  - Is there a difference in the time it takes for an older or younger candidate to become a Mountain Leader? If so, why do you think that is?

- We've spoken about how a candidate's age at registration might influence their chances of completion, now, could you tell me a little bit about how life experience at registration might affect candidates' chances of completion?
  - Prompt: positively or negatively

- Are there are any professions that significantly influence, positively or negatively, completion of the Mountain Leader award?
  - What do you think it is about these professions that makes a difference here?
    * Prompt: Time, money, relatedness, soft skills, dealing with large groups of diverse people (teachers, coaches, youth leaders etc.), doing work in difficult situations (e.g., A&E doc/nurse/teacher)

- Is there anything about candidates' backgrounds that you think is important but we haven't spoken about?
- We have spoken about a number of different factors relating to candidate background. Do you think that there any factors relating to candidate background that are generally more important with regards to completion of the Mountain Leader award?

A.1.2 Candidate Career History and Social Influence.

I would like to find out a little bit more about the Mountain Leader candidates in your region; what has led them to the Mountain Leader award, why they might be interested in it, and what they might do with the award.

A.1.2.1 Motives/Motivation.

- Could you give me some examples of the types of reasons that people have given for wanting to become a Mountain Leader?
  - Prompt: to develop their personal skills/to become an instructor. Taking family & friends out/working, scouting
- How, if at all, do you think that these different motives have affected candidates' chances of becoming a Mountain Leader?
  - Prompt: doing it as a means to an end; doing it because they love being outdoors
  - Prompt: …and what about candidates who are enrolled on an "outdoor degree"/centre trainee/fast-track scheme?

A.1.2.2 Intentions/Expectations.

- Could you give me some examples of different time scales that candidates expect to become Mountain Leaders in?
  - …so, you have mentioned that some people expect to complete quite quickly while others expect it to take a long time; why do you think that there are these differences?
  - How often, if at all, do you see discrepancies between people's expectations and reality?
  - How, if at all, does the time scale a candidate expects to become a Mountain Leader in affect their chances of completion?

We have just been discussing some of the reasons that people want to become Mountain Leaders and how long they think it will take them. Now I would like to find out how hopeful Mountain Leaders view their future career.

- What do candidates think working as a Mountain Leader will be like?
- Which candidates' expectations are the closest to reality?
  - …and whose are the furthest from reality?
- Why do you think these discrepancies exist?
- How do these discrepancies affect a candidate's chances of becoming a Mountain Leader?
- Could you give me some examples of how different candidates might see the Mountain Leader award fitting into their working lives?
  - Long- vs. short- term career?
  - ML only or aspirations to hold higher awards?
  - Wholly- vs. partly- in the mountains?
  - Interactions between the 3
- Some candidates will see themselves working wholly in the mountains once they become Mountain Leaders, whereas others won't could you tell me a little bit about the differences you see in these candidates?
  - Can you give me some examples of the types of people who see themselves as working wholly in the mountains?
    - Which of these candidates do you think expect that working in the mountains will be a long-term career for them?
    - …and which of the see it as a shorter-term career?
    - Do these candidates see the Mountain Leader award as a stepping stone to higher awards (e.g., MIA or IML) or, do they see becoming Mountain Leaders as their ultimate goal?
    - Do you see any differences in the expectations of older and younger candidates with regard to their future careers being wholly or partly in the mountains?
  - …and can you give me some examples of candidates who see themselves as working partly in the mountains and partly elsewhere?
    - Which of these candidates do you think expect that working in the mountains will be a long-term career for them?

        - …and which of the see it as a shorter-term career?
        - Do these candidates aspire to hold higher level awards, or do the only want to become Mountain Leaders?
    - Do you feel that candidates who only want to become Mountain Leaders try to complete the award quickly so that they can start working, or do they take their time as they feel that there is no rush?
    - Do you feel that candidates with aspirations for higher awards present themselves for an assessment before they are ready to try and move along up the qualification ladder, or do they present well prepared in an effort to be efficient with their time?

- Do you think that candidates who see outdoor instruction as being their main source of income and aspire to holding higher level awards are different from those who see out door instruction as their main source of income but do not aspire to hold higher level awards?
    - What do you think the difference is here?
- Are there people who have these higher aspirations, but don't see outdoor instruction as their only source of income?
    - Could you tell me a little bit about these people?
- We have spoken a little bit about the different views and expectations candidates might hold; how, if at all, do you think that these expectations affect a candidate's progression through the Mountain Leader award?
    - Prompt: who become Mountain Leaders? & Time to completion
        - What do you think is different here?
- Have you ever seen a candidate's end goal change as they progress through the Mountain Leader pathway? If so, what changed and why do you think that was?
    - Can you give me some examples of the types of people whose end goal has changed?
        - Have you ever seen the opposite?
    - What effect do you think that these changes have on a candidate's chance of completion? Relevant work experience

- Could you tell me about the opportunities available in your region to candidates for relevant work, paid or unpaid, prior to their ML assessment, if there are any?
  - Do you think any of these opportunities more useful to candidates than others? If so, which ones do you think are the most useful?
  - …and which ones do you think are the least useful?
  - Are any of these opportunities unhelpful to candidates? If so, which ones are these and why do you think that is?
- Could you give me some examples of the types of candidates who make use of these opportunities?

A.1.2.3 Critical Developmental Events.

Sometimes, an event in our lives can change our views or may present us with an opportunity for change. This could be something as extreme as losing one's job or, in a less extreme example, this might be spending a day in the outdoors with family, friends or an inspiring instructor. These events are sometimes called "critical developmental events".

- How many candidates, if any, attend training courses following something that could be described as a "critical developmental event"?
- Some of these events might instigate more permanent motivation than others, could you give me some examples of events that has provided candidates with strong but short motivation?
- …and could you give me some examples where perhaps it has been a factor in enduring motivation for candidates?
- How, if at all, do you think that critical developmental events influence candidates?
  - Positive or negative
  - Same for all candidates?
  - Why do you think that might be?

A.1.2.4 Other Qualifications.

- Some candidates will be registered for other sporting qualifications, both with Mountain Training and other organisations; how, if at all, does being registered for more than one qualification affect a candidate's chances of becoming a Mountain Leader?
- Which other qualifications, if any, seem to have the greatest impact on a candidate's chances of becoming a Mountain Leader?
  - Prompt: positive or negative; make them comfortable in the mountains, managing groups; leaves them with little free time; decision making; leadership
- Some candidates will be working towards non-sporting qualifications e.g., a degree or an NVQ. How, if at all, do you think that working towards a non-sporting qualification influences a candidate's chances of completion?
  - Could you give me some examples of candidates who have benefited from working towards a non-sporting qualification at the same time as working towards their Mountain Leader Award?
  - …and could you give me some examples of candidates who have been adversely affected by working towards a non-sporting qualification at the same time?
- Which influences a candidate's chance of completion more, being registered for another Mountain Training qualification, a qualification in another sport (e.g., mountain biking, paddling or mainstream sports) or a non-sporting qualification?
  - …and why do you think that is?

A.1.2.5 Subjective Norms/Social Influence.

We are all involved in different social groups and have different places in them, creating our own "social norms." I would just like to reiterate that I make no assumptions about what is "good" or "bad" with regards to factors that may or may not influence candidates' completion of the Mountain Leader Award.

- Have you observed that people who come from certain social groups fair better on ML courses than people from other social groups? E.g., candidates whose families or friends regularly engage with the mountains? Candidates from inner cities vs.rural areas? Candidates who are sporty vs.those who are less sporty?
  - What is it about those social groups that you think might make those differences?
- Can you give me some examples of social groups that candidates who attend training but not assessment have come from?
  - One definition of a social group is, "two or more people who interact with one another, share similar characteristics, and collectively have a sense of unity"
  - Could you tell me how, if at all, these social groups influenced candidates?
    - Prompt: positively or negatively?
    - Attraction to outdoor- vs. everyday-life
- Are there any social groups that are noticeably absent from Mountain Leader courses? If so, which ones?
- How important, if at all, is it for a candidate to feel that they have people around them who understand why they want to be a Mountain Leader?
- How, if at all, does the sense of being different from other people influence a candidate's chance of completion?
- How important, if at all, is it for candidates to spend time in the mountains with people from their social groups?

A.1.2.6 Relevant Media Influence.

Candidates will not only be influenced by the people they spend time with but also by the things that they read, see, and hear.

- Could you give me some examples of the types of candidate that have been influenced by the things that they have read, watched or heard?
  - Prompt: positive or negative influence
  - Prompt: UKC/H, Trail magazine, The Professional Mountaineer, MT's website

- Which candidates, if any, do you think are more likely to be influenced than others? Why do you think that is? We are all aware that social media has become an increasingly important and useful part of day-to-day life.
- Please would you tell me a little bit about how candidates use social media in relation to the Mountain Leader award?
- Could you give me some examples of candidates and how they have been helped by social media?
- Could you give me some examples of social media helping candidates?
- …and can you give me some examples of candidates being negatively influenced but social media?
- Which candidates, if any, does social media have the greatest impact on?
- Have you noticed any social media groups that appear to be particularly influential? This can be in a positive or negative way.

A.1.2.7 Role Models

- Role models can be both positive and negative (i.e., I want to be like him/her and I don't want to be like him/her). How important do you feel that role models, positive or negative, are to candidates in your region?
- Could you give me some examples of positive role models and their influence on candidates?
  - Prompt: which candidates, time to completion
- Could you give me some examples of negative role models and their influence on candidates?
  - Prompt: which candidates, time to completion
- Do some candidates identify with role models more easily than others?
  - Do you think that it is something about the candidates or the role models that makes it easier to identify with them?
  - What do you think that is?
- Is there anything about candidates' career history or social influence that you think is important but we haven't spoken about?

- We have spoken about a number of different factors relating to candidate career history and social influence. Do you think that there any factors relating to candidate career history or social influence that are generally more important with regards to completion of the Mountain Leader award?

A.1.3 Personal Characteristics.

A.1.3.1 Attitudes/Outcome Expectations.

Now I would like to talk to you a little bit about attitudes and outcome expectations towards Mountain Leader training courses.

- Could you describe some of the attitudes displayed by candidates towards the training course?
  - How, if at all, do these attitudes change over the week?
    - Why do you think that was?
- Could you give me some examples of the type of candidate who attends a training course but have no intention of booking an assessment?
  - Why do you think that these candidates come on a training course?
  - Do any of these candidates go onto assessment? Why do you think that is?
  - Have you ever seen this change over time? If so, what changed their mind?

And now I would like you to tell me about candidates' attitudes and outcome expectations on assessment courses.

- Could you describe some of the attitudes displayed by candidates towards their assessment?
  - How, if at all, do these attitudes change over the week?
    - Why do you think that was?
- Could you give me some examples of the types of candidate who are confident that they will pass, when they arrive for their assessment?
  - What do you think it is about these people that make them think that they are going to pass?

  - …and how did they do?

- …and, could you give me some examples of the types of candidate who are not confident that they will pass, when they arrived for their assessment?
  - What do you think it is about these people that make them think that they are not going to pass?
  - …and how did they do?
- What do you think is different here between those who do expect to pass and those that don't?
- To what extent do candidates see the assessment as a holistic process?
  - …and what effect does this have on their chances of completion?
- Do you see any differences in the attitudes towards assessment between candidates who have received an exemption from training and those who have been on a training course?
  - Do you think it is obvious on an assessment who has received an exemption from training?

A.1.3.2 Self-efficacy.

- To what extent do candidates in your region feel that they have the necessary resources to become Mountain Leaders?
  - How, if at all, do you think this changes over time?
    - Do you think that is that the same for all candidates?
  - How do you think that the ML process facilitates this belief?
    - …again, do you think that is the same for all candidates?
  - And can you give me any other examples of things that facilitate a candidate's belief that they can become a Mountain Leader?
  - Could you give me some examples of obstacles that prevent candidates believing that they are able to become Mountain Leaders?

A.1.3.3 Mastery Aspirations.

Some candidates will want to be the best that they can be, some will want to be good enough to pass, whilst some will want to be better than the people around them.

- Could you give me some examples of the types of candidate who wants to be the best that they can?
- Could you give me some examples of the types of candidate who just want to be good enough to pass?
- What differences, if any, do you think there are there between those who want to achieve the required standard and those who want to be as good as they can be?
- …and can you give me some examples of the types of candidate who want to be better than the other people around them?
- Do you feel that there are any candidates who see an element of competition in the Mountain Leader award? This could be with other candidates, or with people in their day to day lives.
  - Could you give me some examples of the types of candidate who holds these competitive views?
  - What differences, if any, do you see between these candidates and those who don't see any element of completion in the Mountain Leader award?
  - How, if at all, do you think these differences affect candidates' chances of becoming Mountain Leaders?
- How, if at all, do you think that candidates who compare themselves to other candidates benefit from this?
- …and how, if at all, do you think that this is detrimental?
- What effect, if any, do you think that comparing to others on a candidate's chances of completion?
- How, if at all, do you think that the "overall strength" of candidates on a course will affect individual candidates?
  - Do you think that "strong cohorts" to inspire weaker candidates, or are they more likely to discourage them?
    - What, if any, differences do you see with stronger candidates?

  - …and what about "weak cohorts"? Do you think that they might affect a stronger candidate?

- How else do you think that the other candidates on a course might affect candidates?
  - Prompt: completion?

A.1.3.4 Disconfirmatory Experiences.

A disconfirmatory event is one that causes an individual to question their belief in something, sometimes their own abilities.

- Could you give me some examples of disconfirmatory experiences in your region?
  - Prompt: In relation to people, bad days out on the hill
- How common it is for candidates to have disconfirmatory experiences?
- How are candidates affected by these experiences?
  - Prompt: Do they motivate them to prove the person wrong or does it lower their confidence?
- Which candidates, if any, are better at dealing with these experiences than others?
- How, if at all, do disconfirmatory experiences affect a candidate's chances of completion?

A.1.3.5 Resilience & Robustness.

Candidates will need to deal with stressful events, both in the mountains and also in their normal lives. Disconfirmatory experiences could be an example of a stressful event, but there will be many others.

- Some candidates' self-confidence will be knocked by these stressful events, whilst others won't be. What differences do you see between these two groups?
  - Could you give me some examples of the types of candidate whose self-confidence would not be affected?

  - …and could you give me some example of those whose self-confidence would be knocked? Some of those who have had their self-confidence knocked will be better at "bouncing back" from this than others.

- Could you give me some examples of candidates who are better at bouncing back from stressful events that others?

  - …and can you give me some examples of candidates who are less good at bouncing back from stressful events?
  - What is different here?

- When considering who does and who doesn't become a Mountain Leader, which seems more important: the ability to maintain self-confidence or the ability regain self-confidence if it has been knocked?
- Could you explain to me which affects a candidate's chance of completion more: the number of stressful events that they encounter or their ability to deal with these events?
- Is there anything about candidates' personal characteristics that you think is important but we haven't spoken about?
- We have spoken about a number of different factors relating to candidate's personal characteristics. Do you think that there any factors relating to a candidate's personal characteristics that are generally more important with regards to completion of the Mountain Leader award?

A.1.3.6 End.

- Is there anything in relation to: Candidate Background, Candidate Career History, Social Influence or Personal Characteristics that we have not covered but you think is relevant to completing the Mountain Leader Award?
- Thank you for your time today.
- Next time I would like to talk to you about candidate's personal ability and also the support that candidates in your region receive.

A.2 Interview B

Last time we talked about: Candidate Background, Candidate Career History, Social Influence and the Personal Characteristics of candidates. In this part of the study I would like to ask you some questions about Candidate Ability and Candidate Support. Like last time, this interview should take two to three hours and has a break scheduled into it but stop me at any point if you would like another break.

I would like to reiterate that, I make no assumptions about what is "good" or "bad" with regards to becoming a Mountain Leader, so there are certainly no "right" or "wrong" answers. I would like you to speak freely, whatever you say in this interview is between you and the research team.

Like last time I will be recording the interview, please ask for clarification if anything is not clear.

Is there anything that you would like to go over before we start? Maybe something you would like to recap or something that you have thought of since the last interview?

A.2.1 Candidate Experience & Ability.

Now that I have an idea of who candidates are, where they come from, and what it is that has brought them to you; I would like you to tell me about the ranges of candidates' experiences and how good they are at judging their own abilities.

A.2.1.1 Personal Experience.

Candidates will have a variety of experience in the outdoors. I'm sure that there are some candidates who turn up to their assessment with 40 Quality Mountain Days and I am sure that there are others who turn up with twice or even three times that. As well as varying in depth of experience, candidates also vary in the breadth of their experiences. Some will only have experience that falls within the ML remit and others will have a greater breadth of experience in the mountains (i.e., experience that falls outside of the ML remit, this could be in summer or winter).

- Can you tell me how, if at all, you think a candidate's experience in the mountains affects his/her chances passing their assessment?

  - What about this experience has the greatest influence on a candidate's chances of completion?

- Could you give me some examples of candidates with experience outside the ML remit and how this experience has helped and/or hindered them in becoming Mountain Leaders?
- How, if at all, do you feel that the number of QMDs prior to assessment influences a candidate's chances of completion?
  - How much influence, if any, does the content/quality of candidate's QMDs have on their chances of completion?
  - How, if at all, does the period over which a candidate accrues their QMDs affect their chances of completion?
- Are there any occasions where a candidate's level of experience hasn't been reflected by their result? For example, people with little experience passing or people with lots of experience being deferred or failed?
  - …and do you think that this was a fair reflection of them as a "mountain leader"
  - Why do you think some candidates with little experience are able to become Mountain Leaders whilst some more experienced aren't?
- Some people talk about "tickers", candidates that are "ticking off" the prerequisites. Is this something that you have noticed and if so, could you tell me about these "tickers"?
  - Are "tickers" obvious on Mountain Leader courses?
  - Could you give me some examples of the type of candidate who "tick" off QMDs?
  - Do you think that either "tickers" or "non-tickers" are more likely to complete than one another? Why do you think that is?
  - What differences, if any, do you think there are between "tickers" and "non-tickers"?

A.2.1.2 Personal Competency.

- Could you describe someone who is well prepared for their Mountain Leader assessment?
  - What proportion of candidates do you think are like this?
- In your experience, how good are candidates at judging their own abilities?
  - Could you give me some examples of how a discrepancy in perceived- and actual competency might affect a candidate at assessment?
    - Prompt: positively or negatively i.e., better or worse than they thought
  - If there are discrepancies, are they usually in the same direction across the board, or are some candidates better at one thing than they thought whilst also being worse at something else?
  - Which skills, if any, are discrepancies in perceived and actual ability most common in?
    - Prompt: either direction
  - …and which skills are these discrepancies the most significant in?
  - Do you think that the Mountain Leader assessment process helps candidates accurately gauge their own abilities?
- Do you think that there many candidates in your region who are ready for their assessment but don't feel ready and therefore don't book onto an assessment?
  - Could you give me some examples of these candidates?
  - Why do you think it is that they don't feel ready for their assessment?
  - What, if anything, do these candidates have in common?
    - Prompt: Sex, location, age
- …and now what about the training process? Could you give me some examples of how a discrepancy in perceived competency and actual competency affect a candidate on a training course?
  - Prompt: positively or negatively
- Where do you see these discrepancies more, on training courses or at assessment?

  - Does the Mountain Leader training process helps candidates to accurately gauge their own abilities? If so, how does it do this?

- Is there anything about candidates' experience and ability that you think is important but we haven't spoken about?
- We have spoken about a number of different factors relating to candidate experiences' and ability. Do you think that there any factors relating to a candidate's experience and ability that are generally more important with regards to completion of the Mountain Leader award?

A.2.2 Candidate Support.

In this section I am interested in the support that candidates receive from the people around them. This may be directly related to the Mountain Leader award but could also be in their day to day lives.

A.2.2.1 Training Staff/Centre.

- Could you describe some of the attitudes of course staff and directors toward the Mountain Leader award?
  - Prompt: positive or negative
- To what extent, if at all, do you think that course staff and directors feel candidates will become a part of their community by becoming Mountain Leaders?
  - …and to what extent do candidates feel that they will become part of this community by becoming Mountain Leaders?
  - …and what effect do you think this might have?
- How, if at all, do completion rates vary by provider or individual assessor?
  - What makes their completion rates different?
    * Prompt: coaching, preparation, logbook checking, assessment style
- How, if at all, do completion rates vary by training course staff?
  - …and what do you think it is that makes their completion rates different?

- Can you give me some examples of course staff having a positive influence on candidates?
  - …and can you give me some examples of course staff having a negative influence on candidates?
- How, if at all, do training course debriefs vary by provider?
  - …and what effect do you think these differences might have on candidates?
- Are some providers better at helping candidates create development plans than others? If so, what do they do differently?
  - …and what effect do you think these differences might have on candidates?
- How, if at all, do assessment course reports vary by provider?
  - Do some providers give better feedback than others on course reports?
  - How does the content of this feedback vary?
  - …and what effect do you think these differences might have on candidates?
- Do providers offer to help candidates between training and assessment?
  - Could you give me some examples of providers that do this and the help that they offer?
    - …and what effect do you think that this help has on candidates?
    - Which types of candidate do you think that this help has the biggest effect on?
- How common, if at all, is it for candidates to contact providers between training and assessment?
  - What do they ask providers?

A.2.2.2 Instructor Support.

- Could you tell me a little bit about different course staff's coaching and leadership of candidates during training courses?
  - How, if at all, does this influence a candidate's chances of completion?

- Could you give me some examples of training staff's coaching and/or leadership having positively influenced candidates?
- …and could you give me some examples of training staff's coaching and/or leadership having had a negative impact on candidates?
- Different candidates might respond differently to the same leadership and/or coaching style. Can you give me some examples of how different candidates have responded to the leadership and coaching that they have received from course staff?
  - Prompt: Positive and negative
- Could you tell me a little bit about any training opportunities that are available to providers in your region?
  - Which providers make use of these opportunities?
  - …and which ones don't?
  - What do you think is different here?

A.2.2.3 Mentoring.

- To what extent do you think that there are mentoring opportunities available to candidates in your region?
- What do you think that this mentoring does for candidates?
- Do you think that candidates who have been mentored any more or less likely to become mountain leaders than those who have not been mentored?
- Who mentors candidates in your region?
  - Prompt: Formal & informal
- Would you consider centre assistant or fast track schemes to be mentoring?
  - …and what about candidates studying for an outdoor degree?
- How useful is it for Mountain Leader candidates to have a mentor?
- How, if at all, do candidates in your region benefit from mentoring?
- Can give me some examples of mentoring having had a negative influence on candidates?
  - Prompt: reliance on the mentor

A.2.2.4 Social Support.

The purpose of this next section is to try to find out about the help and support that Mountain Leader candidates may or may not receive in your region. This support can be directly related to the candidates Mountain Leader award but also to the rest of their lives. Candidates may receive many different types of support from a variety of people. I will give you a couple of examples, so that you understand what it is I would like you to tell me about.

Candidates might get support from a qualified instructor, providing feedback on technical skills. They might be part of a peer group that encourages them. They might have someone close to them whom they can turn to if they are feeling low. They might be offered financial support. These are some examples of the things that I would like you to tell me about. I am interested in any positive or negative effects of help and support of Mountain Leader candidates receive in your region.

- Do you understand what it is I would like you to talk to me about? Is there anything that you would like me to go over?
- In general, how important, if at all, do you think it is for candidates to be supported?
  - Do you think that there are any types of candidate who need more support that others? Why do you think that is?
  - …and do you think that there are any types of candidates who don't need any support? And why do you think that is?
- Do you think that candidates are being given advice or guidance about the Mountain Leader award in your region?
  - If a candidate were to look for advice or guidance about the Mountain Leader award, who do you think that they would turn to?
    - Prompt: Qualified instructors, MT, providers, employers, friends?
  - Which candidates, if any, need this support more than others?
    - Prompt: Sex, location, age, economic background
  - Are there any candidates who need more advice or guidance than they are receiving? If so, who do you think they are?

  - How much advice or guidance do successful candidates receive?
    - Is this different for candidates who don't become Mountain Leaders?
  - Are there any candidates who have been adversely affected by the advice or guidance they have received from others? If so, who are they?
    - Where do they get this advice from?

- Could you give me some examples of candidates receiving help with practical matters in your region? E.g., financial assistance, help reducing workloads, help planning, refresher courses?
  - If a candidate were to look for practical help with the Mountain Leader award, who do you think that they would turn to?
    - Prompt: Family, friends, charities, MT, providers, employers?
  - Which candidates, if any, need this support more than other?
    - Prompt: Sex, location, age, economic background
  - Are there any candidates who need more practical help than they are receiving?
  - How much practical help do successful candidates receive?
    - Is this different for candidates who don't become Mountain Leaders?
  - Are there any candidates who have been adversely affected by the practical help that they have received from others? If so, who are they?
    - Where do they get this advice from?

- Could you give me some examples of help that candidates receive in dealing with how they feel about the Mountain Leader award? For example, this could be someone helping a candidate with pre-assessment nerves or general encouragement.
  - If a candidate were to look for esteem support, who do you think that they would turn to?
    - Prompt: Friends, family, mentor
  - Which candidates, if any, need this support more than others?
    - Prompt: Sex, location, age, economic background
  - Are there any candidates who are not receiving enough help in dealing with how they feel about the Mountain Leader award?

  - How much help in dealing with their feelings towards the Mountain Leader award do successful candidates receive?
    - Is this different for candidates who don't become Mountain Leaders?
  - Are there any candidates who have been adversely affected by the help in dealing with their feelings toward the Mountain Leader award that they have received from others?

- Could you give me some examples of help that candidates receive in dealing with personal issues relating to their life and future? For example, this could be someone helping them when they feel low, or someone that they can bounce ideas off.
  - Who do candidates turn to for this support?
    - Prompt: Friends, family, partners
  - Which candidates, if any, need this support more than others?
    - Prompt: Sex, location, age, economic background
  - Are there any candidates that need more help in dealing with personal issue than they are receiving?
  - How much help dealing with personal issues do successful candidates receive?
    - Is this different for candidates who don't become Mountain Leaders?
  - Are there any candidates who have been adversely affected by the help in dealing with personal issues that they have received from others? If so, who are they?
    - Where do they get this advice from?

- Could you tell me about how Mountain Leader award candidates support each other? Or do they prepare in isolation of each other?
  - How much help from other candidates do successful candidates receive?
    - Is this different for candidates who don't become Mountain Leaders?
  - Which candidates does this support affect more than others? Why do you think that is?
  - Are there any issues that candidates are more willing to tackle amongst themselves rather than with Mountain Training or course providers?

        * What sort of things are they and why do you think that is?

- We have spoken about a number of different types of support that is available to candidates, how much of this support do you think that they receive from the MTA?
- Could you tell me how important, if at all, you think it is for candidates to be supported by others as they progress through the Mountain Leader award?
    - Do you think that any of these support types (emotional, esteem, tangible, and informational) as more important than any of the others?
- How important, if at all, do you think that providers feel that it is for candidates to be supported through the Mountain Leader award?
    - Do you think that they see any of these support types (emotional, esteem, tangible, and informational) as more important than any of the others?
- …and how important do you think that candidates feel it is that they are supported through the Mountain Leader award?
- Do you think that there are any groups of candidates who need or want more support than others? Maybe more of one type in particular?
- Do you see any differences in candidates who expect more support in general and candidates who seek out more support?
- Could you tell me a little bit about how, if at all, individual differences might affect the influence of different support types?
    - Examples of types of candidates who are impacted more or less by different types?
- Is there anything about candidate support that you think is important but we haven't spoken about?
- We have spoken about a number of different factors relating to candidate support. Do you think that there any factors relating to candidate support that are generally more important with regards to completion of the Mountain Leader award?

A.2.2.5 End.

- Is there anything in relation to: Candidate Ability or Support that we have not covered but you think is relevant?
- Is there anything unique to your region, that we have not discussed, that perhaps makes it different to the others?
- Is there anything we may have missed that you feel would be important? For example, are there factors outside of the MT framework that would give us some insight into completion rates in your region? Even something that you don't quite understand or can't quite explain, something you'd almost feel might sound silly, but somehow you feel is relevant?
- Thank you for your time today

Appendix B

Developing the Survey Tool

B.1 Introduction

The results of Chapter 2 suggested that becoming a Mountain Leader would be influenced by a variety of constructs across several different domains. The aim of Chapter 3 was to collect data from candidates for these constructs to further investigate which factors were the most important for discriminating those who become Mountain Leaders following a training course from those who do not. Given that we wanted to collect data from candidates for 90 variables, using full-length measures of the relevant constructs would be unreasonable for participants, given that it is not uncommon for measures to use more than five items to measure a single variable. Indeed, we believe that if we did measure each construct of interest with a full-length measure, it would create a survey so long that few candidates would complete it and those that did would likely not be representative of the population.

Therefore, the first aim of the work reported in this appendix was to identify a suitable measure for each construct identified in Chapter 2, which could then be used to identify the most important variables for discriminating candidates who do complete the Mountain Leader qualification from those who do not, both in terms of getting to an assessment and passing their first assessment. The second aim of the work reported in this appendix was to reduce the number of constructs that would be included in the survey tool. As such we were particularly interested in identifying short-form measures as using such measures was most likely to allow us to create a suitably short survey to collect data with.

The development of short-form measures to reduce the burden on participants has been of interest to researchers for over 100 years (Smith et al., 2000). However, the development of short-form measures has attracted some criticism (e.g., Levy, 1968; Smith et al., 2000; Wechsler, 1967). One of the main criticisms of short-form measures has been that "rigorous, valid, comprehensive assessment is crucial for the evaluation and treatment of many psychological problems" (Smith et al., 2000, p 102) and that the time saving afforded by a short-form measure does not warrant the loss of validity associated with measuring a construct with fewer items. When creating, or identifying, a short-form measure one should not assume that the evidence for the validity and reliability of the original measure applies to the short-from, therefore it is important to provide evidence for the reliability and validity of the short-form (Smith et al., 2000). This evidence should include, but is not limited to, reliability of the short-from, shared variance between the full- and short-form measure, content validity/coverage of the construct, and also that the reduction in items offers a meaningful reduction in the time taken for the measure to be completed (Horvath and Röthlin, 2018; Smith et al., 2000).

B.2 Method

B.2.1 Measures

Although using full measures was not a realistic aim in the project, we still felt that the reliability and validity of the indicators that we intended to use would use was paramount. Researchers have suggested a variety of ways in which short-form measures can be developed whilst remaining both reliable and valid. Considering the guidance provided by Smith et al. (2000) and Horvath and Röthlin (2018) along with the aim of this research, we identified items which would be used to collect data from candidates using the process detailed below. The aim of the project was to identify the most important discriminatory variables for identifying candidates who do or do not complete the Mountain Leader qualification using Machine Learning techniques rather than testing the relationships between variables using regression-based techniques or structural equation modelling. Therefore, instead of using full-length measures to collect data for each construct, we used one or two item *indicators* for each construct.

Our preference was to identify existing suitable short-form measure (e.g., the Ten

Item Personality Inventory; Gosling et al., 2003). When this was not possible, but there was an existing measure that we were able to access secondary data for, we used the following steps. Firstly, we checked that existing measure did measure the construct of interest and that there was sufficient evidence for its reliability and validity. Secondly, we identified which items we wanted to retain based on both content validity and factor loadings. It was important that the items retained still provided adequate coverage of the construct. In some instances, this meant retaining an item which had a (relatively) low factor loading, but measured a unique aspect of that construct, as opposed to simply retaining items with high factor loadings (regardless of content validity). This approach necessarily lowered the reliability coefficient for the short-form measure; however, it is important to note that internal consistency is only one aspect of validity.

Once we had identified the items we wished to retain, we fitted a single factor latent variable model for both the full- and short-form measure to the secondary data, using `lavaan` (Rosseel et al., 2020), to estimate factor scores for each participant. These factor scores were then used to calculate a Pearson's correlation coefficient between predicted factor sores for full- and short-form measure as an estimate of shared variance. This method is better than correlating the item sum-scores, as latent variables account for measurement error, thus, reducing the likelihood of receiving an optimistically biased estimate due to error correlation. Shared variance with the full measure was our main concern for this aspect of the study, as if the correlations are high enough then the two measures can be thought of as approximately equal (Smith et al., 2000). Finally, we calculated the composite reliability for the new short-form measure (ω; Fornell and Larcker, 1981). If secondary data were not available but we identified a suitable measure, we chose the best item(s) based on face validity of the items and factor loadings reported in the original paper validating the full measure. Finally, if none of the options above were possible, we developed item(s) within the research team in collaboration with Mountain Training.

Below is a brief description of the measures used for each construct. Where available, results of the latent variable correlations are presented in Table B.1. A full list of the items selected for each construct can be found in Table B.4.

B.2.1.0.1 Personality measures.

B.2.1.0.1.1 Big Five. To measure the "Big-Five" personality traits (openness, conscientiousness, extraversion, agreeableness and emotional stability), we used the Ten Item Personality Inventory (TIPI; Gosling et al., 2003). The TIPI comprises ten pairs of items (e.g., "Critical, quarrelsome"), one positively worded and one negatively worded for each trait. Each item has the same stem, "I see myself as…" Participants were then asked to score each item on a seven-point Likert scale from *Disagree strongly* (1) to *Agree strongly* (7) and sum scores were calculated for each of the five traits.

B.2.1.0.1.2 Resilience and Robustness. We used the Brief Resilience Scale (BRS; Smith et al., 2008) to measure resilience and robustness. There is evidence that the BRS can be used to measure these two factors separately (Hardy et al., 2019). We used two items to measure each factor, participants were asked to score each item on a seven-point Likert scale from *Disagree strongly* (1) to *Agree strongly* (7) and sum scores were calculated for each of the five traits. We used data from Hardy et al. (2019) to identify the best indicators of resilience.

B.2.1.0.1.3 Perfectionism. We used items from three subscales of the Frost Multidimensional Perfectionism Scale (FMPS; Frost et al., 1990) to measure two broad dimensions of perfectionism. We used items from the personal standards subscale to measure perfectionistic striving and items from the concerns over mistakes and doubts about actions subscales to measure perfectionistic concerns. We used two items to measure perfectionistic striving and five items to measure perfectionistic concerns; participants were asked to score each item on a seven-point Likert scale from *Disagree strongly* (1) to *Agree strongly* (7) and sum scores were calculated both factors. We used data from Roberts et al. (2013) to identify the best indicators for each construct. Repeating the analyses from Roberts et al. (2013), we found that medium and large effects were still significant, however small effects were not.

B.2.1.0.1.4 Robustness of confidence. We used three items from the Trait Robustness of Self-Confidence Inventory (TROSCI; Beattie et al., 2011) to measure robustness of confidence. Participants used a nine-point numerical rating scale to score each item from *Strongly disagree* (1) to *Strongly agree* (9) with a mid-point anchor,

Neutral (5). We used data from Beattie et al. (2011) to identify the best indicators of robustness of confidence.

B.2.1.0.2 Socio-demographic. Some socio-demographic data are available on the CMS; however, some are not (e.g., income level, education level). To measure these, we used standard socio-demographic questions (e.g., "What is the highest level of school you had completed or the highest degree you had received when you registered?").

B.2.1.0.3 Self-efficacy Scale. Perceived self-efficacy is domain specific and individuals will have varying levels of self-efficacy beliefs across different domains of their lives, therefore it is important that any measure of perceived self-efficacy is domain specific (Bandura, 1997, 2006).

Mountain Training provide clear documentation about what will be required of candidates during their assessment which includes a candidate handbook and syllabus (Mountain Training UK, 2015a), and a separate skills checklist (Mountain Training UK, 2015b). WH conducted an inductive content analysis (Cho and Lee, 2014) of these documents to identify a list of skills, which a candidate should be able to perform on a Mountain Leader assessment. This list of skills was then discussed with Mountain Training's executive officers (N = 5) who agreed that it provided good coverage of the skills that would be covered on an assessment.

Using the list of skills, we created a self-efficacy scale following Bandura's (2006) guidelines. The resultant scale was then piloted with Mountain Training staff (N = 10) who provided feedback on the items, which was used to refine the scale. The final scale comprised eleven items (e.g., "lead a group effectively in the mountains") rated on a scale of *could not do at all* (0) to *highly certain could do* (100) with a mid-point anchor (*moderately could do*; 50). The items could then be presented to participants three times, each with a different introduction as we wanted to measure efficacy at two points along the pathway and candidates' ideal efficacy levels:

1) Please rate how confident you were that you could do them immediately after your training course.
2) Please rate your degree of confidence, as of now/at your (first) assessment.[1]

[1]Different wording was presented to candidates based on whether or not they had been assessed.

3) Now we know about your levels of confidence to perform these tasks as of now/at your (first) assessment, we would like to understand how confident you feel that your ideal self would be/have been at your (first) assessment. The Ideal Self: "Your ideal self is the kind of person you'd really like to be. It is defined by the characteristics you would ideally like to have. It's not necessary that you have these characteristics now, only that you believe you want to have them."

B.2.1.0.4 Personal Projects. We used a modified version of Little's Personal Project Analysis (PPA; Little, 1983), similar to that used by Beattie et al. (2015). We adapted the instructions so that they read:

> We are interested in studying the kinds of personal projects that candidates have at different stages of their life and how they relate to candidates' motivation to become an ML. All of us have a number of personal projects at any given time that we think about, plan for, and sometimes (though not always) complete.
>
> Please take a moment to think about the projects or goals that you were working on before your assessment, these may include things that you have already told us about.

Participants were then given examples of goals (e.g., "Completing another outdoor qualification," "Spending more time with my family") and asked to "write down the two goals that you were most likely to work towards in the six-months before your assessment, *not-including* becoming an ML." On the following page, for each of their stated goals and for the goal of "becoming an ML," they were then asked to rate the: importance of the goal, *not at all important to me* (0) to *extremely important to me* (100); progress towards the goal in the last six months/six months before their assessment, *no progress* (0) to *most progress* (100); and their perceived self-efficacy of attaining the goal, *I definitely do/did not have the skills and resources to be successful at achieving this goal* (0) to *I definitely have/had the skills and resources to be successful at achieving this goal* (100). Using the scores provided, the following can then be calculated: relative importance, relative progress, and relative efficacy score using Equation (B.1).[2]

[2]To avoid returning an undefined value one is added to both the numerator and denominator.

$$Relative = \frac{Mountain\,Leader + 1}{(Goal\,1 + Goal\,2) \div 2 + 1} \tag{B.1}$$

B.2.1.0.5 Motives. In Chapter 2 it appeared that two different levels of motive were important to the completion of the Mountain Leader qualification: participatory (the goal content) and regulatory (the "why"). To measure the participatory motives, we employed a similar methodology to Sheldon and Elliot's (1999) adaptation of Little's (1983) Personal Project Analysis. First, we asked participants to list two goals that they hoped to achieve by registering for the Mountain Leader qualification. These reasons were then coded qualitatively by WH on a scale of *definitely intrinsic* (1) to *definitely extrinsic* (5) and a mean score was calculated. When coding the data, WH was blinded to all outcome variables. Examples of data coded at each each value are: (1) "To have fun," (2) "Being better equipped to enjoy the mountains safely for myself," (3) "Assessing my own ability," (4) "Confidence in leading groups in the mountains," (5) "Gain the ML qualification."

To measure regulatory motives, participants rated each participatory motive they had given in terms of their behavioural regulation. Each item had the same stem, "I pursue this goal because…" The intrinsic item was "of the fun and enjoyment it provides me," the integrated reason was "it is a part of who I am or aspire to be," the identified reason was "I really believe it's an important goal to have," the introjected reason was "I would feel ashamed, guilty, or anxious if I didn't," the external reason was "someone else wants me to or because the situation demands it." Participants scored each of these reasons on a visual analogue scale with five equally spaced anchors from *strongly agree* (0) to *strongly disagree* (100), a mean score for each of the regulatory motives was then calculated.

B.2.1.0.6 Course Staff Coaching Behaviours. The Military Coaching Behaviour Scale (MCBS; Wagstaff et al., 2018) is a 22-item scale that assesses five coaching behaviours: observing and performance analysis, effective questioning, goal setting, developmental feedback, and motivational feedback. We used two items for each factor, scored on a Likert scale from *Not at all* (1) to *All of the time* (7) and sum scores were calculated each factor. We used unpublished data to identify the best indicators of each coaching behaviour.

B.2.1.0.7 Need Supportive Environment. The Perceived Environmental Supportiveness Scale (PESS; Markland and Tobin, 2010) measures autonomy support, structure, and involvement, each with five items. We used one item for each factor, scored on a numerical rating scale from *Not true for me* (0) to *Very true for me* (6). We were unable to obtain data collected using this measure, therefore we chose one item for each factor based on face validity.

B.2.1.0.8 Perceived Conflict on Courses. We used items from the Intra-group Conflict Scale for Sport (ICSS; Boulter et al., 3001) to measure perceived intragroup conflict on courses. In the ICS-S, five items measure relationship conflict, four measure process conflict, and four measure task conflict. We did not measure task conflict in this study as there was no evidence in the qualitative study that it was relevant to completion of the Mountain Leader qualification. We used one relationship conflict item and one process conflict item. Each item was scored on a Likert scale from *None/Never* (1) to *A lot/Always* (9). We asked each of the items in the context of conflict between candidates and between candidates and staff, four items in total. We used data from Boulter et al. (3001) to identify the best indicators of relationship and process conflict.

B.2.1.0.9 Social Support. We considered four dimensions of social support (i.e., esteem, emotional, informational, and tangible support) in two contexts, perceived available support and received support. We used two items from each dimension of The Perceived Available Support in Sport Questionnaire (PASSQ; Freeman et al., 2011) to measure perceived available support and two items from The Athletes' Received Support Questionnaire (ARSQ; Freeman et al., 2014) to measure received support. All items were scored using a Likert scale, with the options *Not at all* (1) to *Extremely so* (5) for the PASSQ items and *Not at all* (1) to *Seven or more times* (5) for the ARSQ items. We used data collected as part of development of the ARSQ (Freeman et al., 2014) to identify the best indicators of perceived and received support.

B.2.1.0.10 Preparation for Assessment. Preparation for an assessment may encompass a variety of different things for different candidates and we were interested in how much candidates felt that they had done to prepare for an assessment. We asked participants to complete the sentence, "I have done _____ to prepare effectively for an

ML assessment" using a visual analogue scale, anchored at *nothing* (0) and *all that I could* (100). Given the complex nature of this question, we used a decomposition approach (cf. the World Health Organization Health and Work Performance Questionnaire Kessler, 2003; Means and Loftus, 1991) to improve accuracy in responses by first asking participants to list some of the things that they had done in the last six-months/six-months prior to their assessment to prepare. The aim of this approach is to bring relevant activities to mind, so that when participants complete the question of relevance, they are able to do so more accurately (cf. Kessler, 2003).

B.2.1.0.11 Life events. Based on the results of Chapter 2, we wanted to measure change in three domains of candidates lives: social, professional, and health. The Recent Life Change Questionnaire (RLCQ; Miller and Rahe, 1997) has items covering these domains. At this point, we were not concerned about the exact events that may, or may not, have occurred. Therefore, we presented items from the RLCQ as examples for each domain. Participants were asked to rate the extent to which they had experienced change in that domain of their life since their training course using a visual analogue scale from *No change* (0) to *Major change* (100). Another consideration when choosing this method was the sensitive nature of some life events. Allowing participants to indicate a magnitude of perceived change rather than explicitly responding to a sensitive item (e.g., "Miscarriage or abortion," "Being held in jail") was deemed more appropriate for this study.

B.2.1.0.12 Aspirations, Intentions, and Expectations. To understand what candidates hoped to achieve, their intentions towards assessment, and how long after their training course candidates thought that they would be assessed, if they intended to do so, we created items in conjunction with Mountain Training as no measures existed.

B.2.2 Participants and Procedure

We created four surveys, each of which contained a subset of the variables that we wanted to collect data for. Each variable was included in at least two of the surveys and each pairwise combination of variables was included in at least one survey. This was done to both collect as much data as possible and to ensure that two-way interactions

Table B.1: Latent variable correlations between full- and short-form measures.

Measure	Variable	n	r	95% CI	ω_{full}	ω_{short}
FMPS	Perfectionistic strivings	120	.81*	.74,.87	.80	.52
	Perfectionistic concerns		.75*	.66,.82	.74	.51
BRS	Resilience	192	.97*	.96,.98	.91	.83
TROSCI	Robustness of confidence	267	.89*	.87,.91	.81	.68
MCBS	Observation	263	.96*	.95,.97	.96	.86
	Effective questioning		.93*	.92,.95	.96	.79
	Goal setting		.95*	.93,.96	.96	.80
	Developmental feedback		.95*	.94,.96	.98	.83
	Motivational feedback		.97*	.97,.98	.98	.87
ICSS	Relationship	384	.80*	.76,.83	.88	NA
	Process		.79*	.74,.82	.85	NA
PASSQ	Emotional	219	.97*	.96,.97	.90	.85
	Esteem		.95*	.93,.96	.85	.78
	Informational		.90*	.87,.92	.82	.74
	Tangible		.83*	.78,.86	.82	.67
ARSQ	Emotional		.95*	.94,.96	.91	.86
	Esteem		.95*	.94,.96	.91	.84
	Informational		.94*	.92,.95	.91	.76
	Tangible		.95*	.93,.96	.92	.83

Note:
ω can only be calculated when there is more than one item.
* $p < .01$

between variables could be explored. Figure B.1 shows a simplified visual representation of the distribution of constructs between the four surveys.

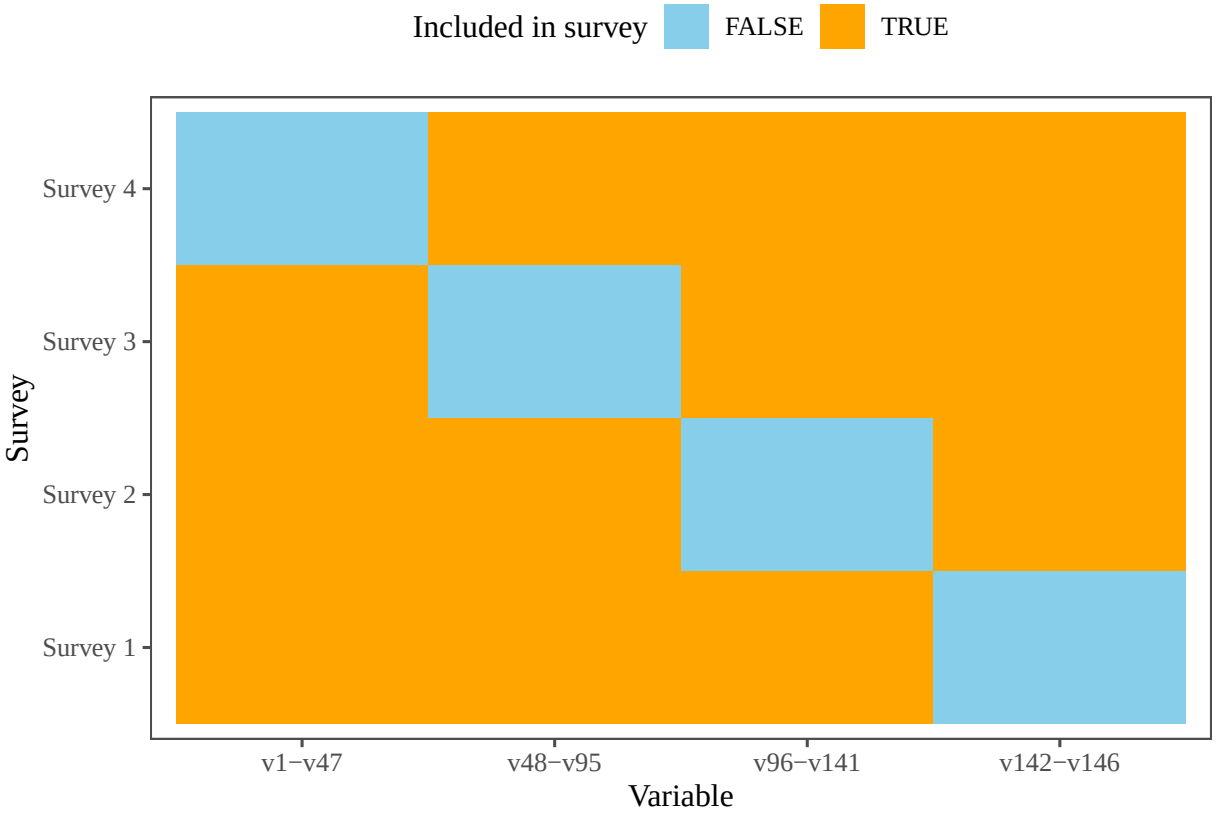

Figure B.1: Simplified representation of variable overlap between Groups 1 to 4.

In November 2018, we contacted all candidates trained between 2008 and 2016 ($n = 3794$). None of these candidates were included in the main analyses for Chapter 3, nor were they included in the item selection work reported above in the Appendix. Each candidate had been randomly assigned to one of four groups (stratified by year of training) using the `randomizr` package (Coppock, 2019) and candidates from each of these groups were invited to complete one of the surveys described above using the Qualtrics survey platform (Qualtrics, 2019). We collected responses from 1056 participants (27.83 % response rate), see Table B.2 for summary statistics of participant demographics within each group.

Once data collection was complete, each of the four groups was then split in two, one group for those candidates who had DLOG data and one group who did not have DLOG data. This was done as the pattern recognition procedure cannot handle missing data and we would then have had to omit all DLOG data, which would have left us unable to identify interactions between the survey and experience data. Once the groups

Table B.2: Survey participants per group

Group	n	Female (%)	M_{Age}	$M_{years\,since\,training} \pm 1$ SD
1	260	23.46	38.31	5.65 ± 2.58
2	264	27.65	37.24	5.72 ± 2.61
3	266	19.92	39.93	5.61 ± 2.55
4	266	25.94	38.25	5.71 ± 2.55

had been split into these two groups, we created two data sets within each one for those who did not and then for each classification problem. This process resulted in the following data sets for each survey group (Figure B.2 provides a visual representation of the groups described below):

1) Getting to assessment within 18 months of training - no DLOG data.
2) Getting to assessment within 18 months of training - with DLOG data.
3) Passing the first assessment - no DLOG data.
4) Passing the first assessment - with DLOG data.

In our data (and the population), most candidates have not been assessed 18 months after their training course. To ensure an orthogonal design (i.e., outcome groups of equal size) we selected a random sample of candidates who had not been assessed 18 months after their training course of equal size to the group of candidates who had been assessed.

> (Could/should I do something to check the representativeness of the samples (e.g., using sex, age, board)? Equivalence testing?)

B.2.3 Analytical Method

We used the same pattern recognition procedure as described for the main study with the following modification. The data were standardised within sex, using the `mousetrap` package (Kieslich et al., 2019), to control for sex differences. This was done as not all groups would have had enough data from both sexes to analyse separately having already split the data by DLOG availability. We considered exploring the data having standardised them by training provider, however, when grouped by training provider, there were not enough cases in each group to standardise the data in a meaningful way.

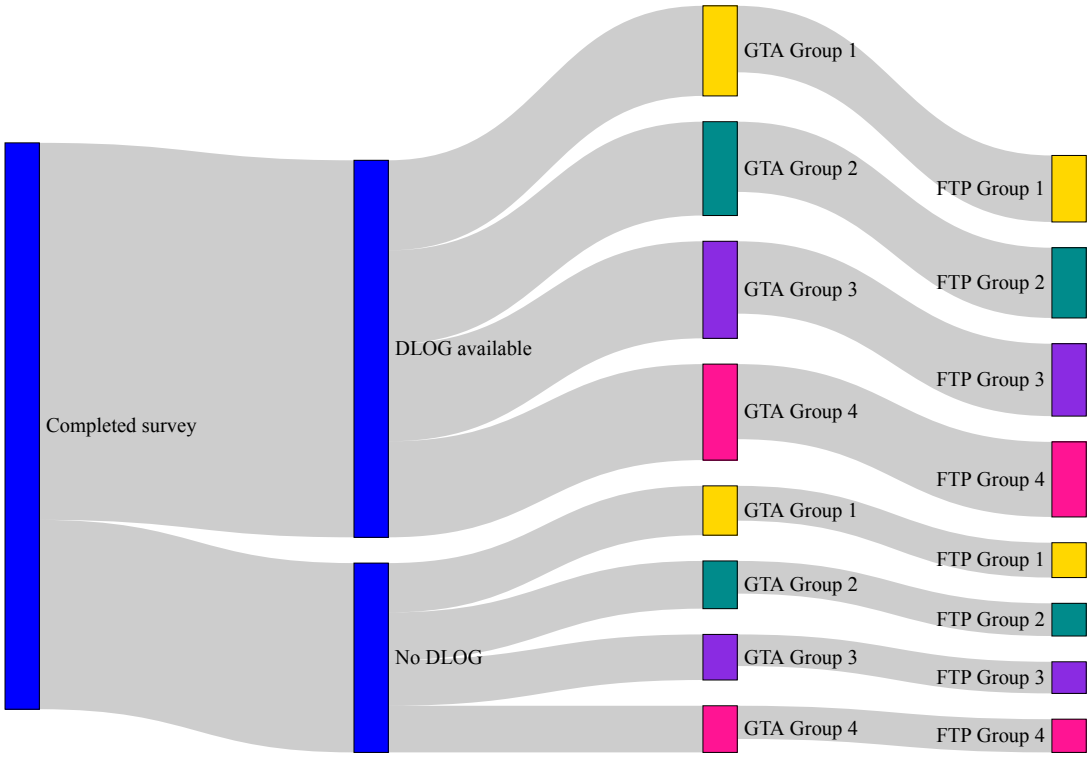

Figure B.2: Study 2 participants split into 16 data sets for analysis. Note: DLOG = Digital logbook, GTA = Getting to assessment within 18 months of training, FTP = Passing the first assessment.

Individual items *and* construct sum-scores were included in the analyses to identify if it was specific elements of a construct that was important, or if it was the construct as a whole.

B.2.4 Item Retention

Having identified a number of feature subsets that could be used to classify candidates in each group, we identified the items which we wanted to retain for the final survey. As not all items were asked to the same number of groups, we scored each item by the number of times that it was selected divided by the number of times it was asked. This was done so that the item retention process was not biased by the number of times that an item was asked. Items were retained if they were selected for the best models in at least half of the datasets they were asked to.

B.3 Results

B.3.1 Item Reduction

Using the process described above for identifying suitable short-form measures, 198 items were removed from full-measures, leaving an item pool of 184 items. Assuming seven seconds per item (Qualtrics, 2019), this equates to a survey that would require candidates to spend approximately 21.47 minutes answering questions (23 minutes shorter than using full-measures); participants would also be required to read the information sheet, transition between pages, etc. The items for 11 variables had evidence of validity from other studies, 61 had evidence for validity from analyses carried out in this study on secondary data, 59 were self-efficacy items created specifically for this project, nine non-DLOG variables were collected from CMS, and 11 variables were sum scores.

We retained 66 variables based on the criteria above. Some of these variables were sum totals of constructs including variables not selected, therefore we chose to retain a further 23 items. This process resulted in 134 items being retained for the final survey, which we estimated would take 12 minutes to complete (see supplementary information for a list of the variables retained for the final survey).

Table B.3: Classification rates for the feature subset with the highest classification rates for each data set (percentage accuracy).

				Classifier percentage accuracy			
Analysis	Group	DLOG	Subset	J48	IBk	NB	SMO
Getting to assessment within 18 months of training	1	TRUE	3s	74.49	60.20	70.41	73.47
		FALSE	3s RFE	79.17	76.39	77.78	76.39
	2	TRUE	4s	75.51	76.53	82.65	78.57
		FALSE	2s RFE	79.49	79.49	80.77	87.18
	3	TRUE	2s	66.67	61.90	65.87	68.25
		FALSE	2s RFE	54.55	66.67	74.24	80.30
	4	TRUE	3s	71.19	70.34	78.81	77.97
		FALSE	3s	65.38	71.79	69.23	73.08
Passing first time	1	TRUE	2s RFE	61.54	67.31	76.92	75.00
		FALSE	3s	63.89	63.89	61.11	75.00
	2	TRUE	2s RFE	81.25	81.25	84.38	87.50
		FALSE	3s	50.00	73.81	64.29	57.14
	3	TRUE	3s	69.44	72.22	77.78	83.33
		FALSE	2s	72.50	45.00	62.50	67.50
	4	TRUE	3s	69.44	72.22	77.78	83.33
		FALSE	2s	72.50	45.00	62.50	67.50

B.3.2 Classification Rates

In 15 of the 16 datasets we were able identify a feature subset which could be used to correctly classify candidates with at least *good* accuracy (i.e., at least one classifier for that data set had a classification rate over 70%). For the other data set we were able to identify a feature subset which could be used to classify candidates with *moderate* accuracy (i.e., at least one classifier for that data set had a classification rate over 60%). Whilst these classification rates are not as high as one may like, we believe that they are acceptable because no survey contained all all of the variables that we considered to be potentially important. Table B.3 shows the classification rates for the best models within each data set.

B.4 Discussion

This study sought to create a survey tool which could be administered to candidates who had completed a Mountain Leader training course in order to help us identify the most important discriminatory variables for candidates who: (a) did or did not get to an assessment within 18 months of their training course and (b) did or did not pass their first assessment. Whilst no single candidate provided data for all the variables, we were able to discriminate candidates with a degree of accuracy substantially greater than chance in each group. This finding shows that firstly, the measures used in the survey work and secondly, that we collected data about variables which explain some of the variance in the criterion variables. It is important to note that just because a construct has not been selected, it is not necessarily unimportant. Variables not selected as discriminatory variables may in fact be important commonalities between the groups. The next study collected data from candidates who attended a Mountain Leader training course between 2016 and 2018 on all the variables retained following this study.

Including DLOG data in the models did not appear to improve the classification rates in any substantive way. This finding suggests that the variance explained by those data is better explained by survey variables. A likely explanation is that candidates use the DLOGs in different ways. Some candidates will log every experience that they have, some will log only the best of their experiences, some will log only their relevant experience, and some will log only the experience they need to meet the prerequisites for the course (potentially from an extremely large pool of experience). The use of DLOG in these different ways creates "messy" data, with no easy way to distinguish a candidate who only has 40 QMDs and a candidate who has far more than that but only logs 40 as they do not feel it would benefit them to log more.

B.5 Study 2 Supplementary Information

Table B.4: Survey variables.

Variable	Measure	Item wording	Number of items	
			Full	Short
Perfectionistic strivings	FMPS	1) It is important to me that I be thoroughly competent in everything I do 2) I set higher goals than most people	7	2
Perfectionistic concerns	FMPS	1) If I fail partly, it is as bad as being a complete failure 2) People will probably think less of me if I make a mistake 3) The fewer mistakes I make, the more people will like me 4) Even when I do something carefully, I often feel that it is not quite right 5) I usually have doubts about the simple everyday things I do	9	5
*Resilience	BRS	1) I tend to bounce back quickly after hard times. 2) I tend to take a long time to get over set-backs in my life. (R)	4	2
*Robustness	BRS	1) I have a hard time making it through stressful events. (R) 2) I usually come through difficult times with little trouble.	2	2
Robustness of confidence	TROSCI	1) Negative feedback from others does not affect my level of self-confidence 2) Mistakes have very little effect on my self-confidence 3) My self-confidence is stable; it does not vary much at all	9	3
*Extraversion	TIPI	I see myself as… 1) Extraverted, enthusiastic. 2) Reserved, quiet. (R)	2	2
*Agreeableness	TIPI	I see myself as… 1) Critical, quarrelsome. (R) 2) Sympathetic, warm.	2	2
*Conscientiousness	TIPI	I see myself as… 1) Dependable, self-disciplined. 2) Disorganised, careless. (R)	2	2
*Emotional stability	TIPI	I see myself as… 1) Anxious, easily upset. (R) 2) Calm, emotionally stable.	2	2
*Openness to experience	TIPI	I see myself as… 1) Open to new experiences, complex. 2) Conventional, uncreative. (R)	2	2
Highest education level at registration	SSHES	What is the highest level of school you have completed or the highest degree you have received?	1	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
Income level	SSHES	Information about income is very important to understand. Would you please give your best guess? Please indicate the answer that included your entire household income in (previous year) before taxes.	1	1
*Intention to complete at registration	SSHES	When you first registered for the Mountain Leader qualification, did you intent to continue to assessment?	1	1
*Expected time to completion at registration	SSHES	When you registered for the Mountain Leader qualification, how many months after your training course did you intend to be assessed?	1	1
Holiday entitlement (paid) at registration	SSHES	How many days paid holiday, if any, do you think you were entitled to a year at that point? Don't worry if you are unsure, please just use the most accurate number you think you were entitled to.	1	1
*Perceived travel time to the nearest mountainous region at registration	SSHES	How long do you think it would have taken for you to travel from your home at the time [of registration] to the nearest mountainous area? (To the nearest half hour)	1	1
*Perceived ease access to the mountains at registration	SSHES	Using the slider below, how easy do you feel that it was for you to get to the nearest mountainous area?	1	1
Aspirations at registration	SSHES	Candidates who have registered for the ML may have different aspirations. Below is a list of common aspirations. Please tick the option which best reflected your aspirations at registration, you may only choose one.	1	1
*Registered for multiple qualifications at registration for ML	SSHES	Some candidates are registered for other qualifications as well when they register for the Mountain Leader qualification. Please select the option that best suits you: When I registered for the Mountain Leader qualification I was registered for…	1	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Participatory motives for registering	PPA	People have different reasons/goals when they register for the Mountain Leader qualification. Please write two goals you hoped to achieve by registering for the Mountain Leader qualification.	5	2
*Level of intrinsic motivation for registering	PPA	I pursued this goal because of the fun and enjoyment it provided me.	5	2
Level of integrated motivation for registering	PPA	I pursued this goal because it as part of who I was or aspired to be.	5	2
Level of identified motivation for registering	PPA	I pursued this goal because I really believed it was an important goal to have.	5	2
*Level of introjected motivation for registering	PPA	I pursued this goal because I would have felt ashamed, guilty, or anxious if I didn't.	5	2
*Level of external motivation for registering	PPA	I pursued this goal because someone else wanted me to or because the situation demanded it.	5	2
*Importance of becoming an ML	PPA	Assessed: Using the sliders below, please rate how important each goal was to you in the six-months prior to being assessed.	6	3
*Progress towards becoming an ML	PPA	Assessed: Using the sliders below, please rate how much progress you made with this goal in the six-months prior to being assessed.	6	3
*Efficacy to become an ML	PPA	Assessed: Using the sliders below, please rate to what extent you felt that you had the skills and resources to be successful at achieving this goal (as a percentage).	6	3

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Perceived available time to become an ML	SSHES	We have asked about a number of different aspects of your life, including other things you did in your life (personal projects, profession, etc.). Thinking about your life as a whole, to what extent did you feel that you had enough available time to become a Mountain Leader	1	1
*Understanding of the qualification when completing the survey	SSHES	As a percentage, how confident are you now in your understanding of the purpose of the Mountain Leader qualification?	1	2
*Recalled understanding of the qualification pre-training	SSHES	As a percentage, how confident were you before your training course in your understanding of the purpose of the Mountain Leader qualification?	1	2
*Intention to complete at the start of training	SSHES	When you arrived for your training course, did you intend to continue to assessment?	1	1
*Expected time to assessment at start of training	SSHES	When you arrived for your training course, how many months after your training course did you intend to be assessed?	1	1
*Perception of training course staffs' "observation" skills	MCBS	The staff on my training course… 1) Paid close attention to what I did. 2) Carefully observed my skills.	4	2
*Perception of training course staffs' "effective questioning" skills	MCBS	The staff on my training course… 1) Encouraged me to question the way I did things. 2) Encouraged me to make suggestions on how I could improve my performance.	4	2

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Perception of training course staffs' "goal setting" skills	MCBS	The staff on my training course… 1) Helped me identify targets for attaining my goals. 2) Helped me set long term goals.	4	2
Perception of training course staffs' "developmental feedback" skills	MCBS	The staff on my training course… 1) Made sure I understood what I needed to do to improve. 2) Gave me advice on how to improve my skills.	4	2
*Perception of training course staffs' "motivational feedback" skills	MCBS	The staff on my training course… 1) Expressed appreciation when I performed well. 2) Told me when I did a particularly good job.	4	2
*Perceived provision of autonomy by training staff	PESS	During my training course, the staff provided me with choices and options.	5	1
*Perceived provision of structure by training staff	PESS	During my training course, the staff made it clear to me what I need to do to get results.	5	1
*Perceived involvement of training staff	PESS	During my training course, the staff made enough time for me even though they were busy.	5	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Perception that the training course felt like an assessment	SSHES	To what extent do you agree with the following statement: "On my training course, I felt like I was being assessed constantly."	1	1
Type of post-training debrief	SSHES	Which of the following statements best describes your post-training debrief?	1	1
Expected change in Expected time (in months) to assessment post-debrief	SSHES	Please place the slider at the point you feel best completes this sentence for you: "Following the feedback from my training course, I felt that it would take _____ I had previously thought to prepare for an assessment."	1	1
Perceived understanding of how to prepare effectively for an assessment	SSHES	Following your debrief, how well did you feel that you knew what you needed to do to prepare for an assessment? "I _____ what I needed to do to prepare effectively for an assessment."	1	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Post-training self-efficacy	SSHES	The attached form lists different activities that are involved in the Mountain Leader qualification. We are interested in how confident you were that you could carry out the following actions following your training course. Please rate how confident you were that you could do them immediately after your training course. Rate your degree of confidence by dragging the bar to record a number from 0 to 100 using the scale given below: 1) Wild camp for two nights in any weather. 2) Choose appropriate routes whilst leading others in the mountains. 3) Choose appropriate equipment for mountain walking and explain the choice. 4) Look after myself and other in steep ground/crossing a river. 5) Lead a group effectively in the mountains. 6) Provide immediate medical care in the mountains. 7) Navigate to a chosen point on a map in any conditions, night or day. 8) Plan a mountain day that is appropriate for the group. 9) Respond appropriately to an emergency (e.g., a broken leg). 10) Act according to my responsibilities to others (e.g., group members, parents/guardians, employers). 11) Look after the mountain environment and encourage others to do so too.	11	11
*Perceived level of challenge on the training course	SSHES	To what extent do you agree with the following statement? "I found the training course too challenging."	1	1
*Perceived relationship conflict between candidates on the training course	ICSS	For these statements, please think about the interactions between candidates (including yourself) on your training course. How much personality conflict was evident on your course?	5	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
Perceived process conflict between candidates on the training course	ICSS	For these statements, please think about the interactions between candidates (including yourself) on your training course. To what extent did you disagree about the way to do things on your course?	4	1
*Perceived relationship conflict between staff and candidates on the training course	ICSS	Now we would like you to think about the interactions between candidates and the staff on your training course. How much personality conflict was evident on your course?	5	1
Perceived process conflict between staff and candidates on the training course	ICSS	Now we would like you to think about the interactions between candidates and the staff on your training course. To what extent did you disagree about the way to do things on your course?	4	1
Financial support for training course	SSHES	We are interested in understanding how, if at all, financial support influences progress through the Mountain Leader scheme. Please use the slider below to indicate what percentage of your training course fee you paid yourself:	1	1
Change in understanding of the purpose of the ML qualification post-training	SSHES	What influence, if any, did the training course have on your understanding of the purpose of Mountain Leader qualification?	2	2

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Change in perception of the standard of the ML post-training	SSHES	Some people talk about "the standard" required of candidates in order for them to become Mountain Leaders. For some people, the training course changes their view of the standard required to become a Mountain Leader, whereas for others, it confirms what they previously believed the standard to be. Following your training course, where did you feel the standard was in comparison to your perception of the standard prior to your training course? The standard was _____ I thought it was before the training course	1	1
*Intention to complete at the end of training	SSHES	By the end of your training course, did you intend to continue to assessment? I had _____ of being assessed	1	1
*Expected time to assessment at the end of the training course	SSHES	By the end of your training course, how many months did you think it would be until you were assessed?	1	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Experience of change post-training	RLCQ	There are a number of events in all our lives that will influence other aspects of them. We would like to understand how events in different parts of your life may or may not have influenced your progression through the Mountain Leader qualification. Below is a list of different areas of peoples' lives with some examples of events they may experience within them. Social: Children moving out from home, Gaining or losing an immediate family member (adoption, birth, death), Marriage/divorce, Moving to a new home, Becoming a carer for a relative/friend. Professional: Changing job, Increased/decreased income, Retirement, Change in working hours. Health: Illness or injury requiring medial attention, Injury that reduces mobility, Back ache/muscular pain.	66	1
*Negative experiences post-training	SSHES	Some candidates have less than positive experiences following their training course. Please indicate when, if ever, you experienced any of the following: 1) Getting lost in the mountains unexpectedly 2) Negative comments from others online, relating to you and your aspirations to become a Mountain Leader 3) Negative comments from others in person, relating to you and your aspirations to become a Mountain Leader	18	18
*Attended additional formal training post-training course	SSHES	Have you attended any additional formal (i.e., with an instructor) training following your (first) Mountain Leader training course?	1	1
*Perceived availability of emotional support	PASSQ	If needed, to what extent would someone have… 1) always been there for you 2) shown concern for you	4	2

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Perceived availability of esteem support	PASSQ	If needed, to what extent would someone have… 1) instilled you with the confidence to deal with pressure 2) boosted your sense of competence	4	2
*Perceived availability of informational support	PASSQ	If needed, to what extent would someone have… 1) given you advice about becoming a Mountain Leader 2) given you advice about performing at assessment	4	2
*Perceived availability of tangible support	PASSQ	If needed, to what extent would someone have… 1) helped you with tasks to leave you free to concentrate on becoming a Mountain Leader 2) helped you organise and plan your preparation/consolidation	4	2
*Actual received emotional support (in the last week/week before assessment)	ARSQ	Not assessed: In the last week, how often did someone… Assessed: In the week before your assessment, how often did someone… 1) show concern for you 2) make you feel that they would always be there for you	4	2
*Actual received esteem support (in the last week/week before assessment)	ARSQ	Not assessed: In the last week, how often did someone… Assessed: In the week before your assessment, how often did someone… 1) tell you, you can do it 2) boost your confidence	4	2
*Actual received informational support (in the last week/week before assessment)	ARSQ	Not assessed: In the last week, how often did someone… Assessed: In the week before your assessment, how often did someone… 1) offer you ideas and suggest actions 2) give you advice about what to do	4	2

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Actual received tangible support (in the last week/week before assessment)	ARSQ	Not assessed: In the last week, how often did someone… Assessed: In the week before your assessment, how often did someone… 1) help plan your preparation/consolidation 2) help you with tasks	4	2
*Financial support for assessment course	SSHES	Some people receive financial support for their Mountain Leader assessment and others pay for it themselves. Please use the slider below to indicate what percentage of your (first) assessment course fee you paid yourself:	1	1
Perceived affordability of becoming a Mountain Leader	SSHES	To what extent do you agree with the following statement: "I feel that I can afford to become a Mountain Leader."	1	1
*Importance of becoming a Mountain Leader to others	SSHES	We are interested in how important other people in your life (e.g., friends, family, employers) feel it is that you become a Mountain Leader. Please rate how important other people in your life feel it is that you become a Mountain Leader:	1	1
Main source of social support	SSHES	Which of the following sources do you feel that you have received the most support from in your efforts to become a Mountain Leader?	1	1
*Perceived preparation in the last six months/six-months before assessment	SSHES	Overall, how much preparation for a Mountain Leader assessment do you feel that you have done in the last six months? "I have done ____ to prepare effectively for a Mountain Leader assessment	2	2

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Pre-assessment self-efficacy	SSHES	Not assessed: Please rate your degree of confidence, as of now, by dragging the bar to record a number from 0 to 100 using the scale given below. Assessed: Please rate your degree of confidence, at your (first) assessment, by dragging the bar to record a number from 0 to 100 using the scale given below: 1) Wild camp for two nights in any weather. 2) Choose appropriate routes whilst leading others in the mountains. 3) Choose appropriate equipment for mountain walking and explain the choice. 4) Look after myself and other in steep ground/crossing a river. 5) Lead a group effectively in the mountains. 6) Provide immediate medical care in the mountains. 7) Navigate to a chosen point on a map in any conditions, night or day. 8) Plan a mountain day that is appropriate for the group. 9) Respond appropriately to an emergency (e.g., a broken leg). 10) Act according to my responsibilities to others (e.g., group members, parents/guardians, employers). 11) Look after the mountain environment and encourage others to do so too.	11	11

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Ideal pre-assessment self-efficacy	SSHES	Not assessed: Now we know about your levels of confidence to perform these tasks as of now, we would like to understand how confident you feel that your ideal self would be at assessment. Please rate how confident you feel your ideal self would be at a Mountain Leader assessment. Rate your degree of confidence by dragging the bar to record a number from 0 to 100 using the scale given below: Assessed: Now we know about your levels of confidence to perform these tasks at your (first) assessment, we would like to understand how confident you feel that your ideal self would have been at your (first) assessment. Please rate how confident you feel your ideal self would have been at your (first) Mountain Leader assessment. Rate your degree of confidence by dragging the bar to record a number from 0 to 100 using the scale given below: 1) Wild camp for two nights in any weather. 2) Choose appropriate routes whilst leading others in the mountains. 3) Choose appropriate equipment for mountain walking and explain the choice. 4) Look after myself and other in steep ground/crossing a river. 5) Lead a group effectively in the mountains. 6) Provide immediate medical care in the mountains. 7) Navigate to a chosen point on a map in any conditions, night or day. 8) Plan a mountain day that is appropriate for the group. 9) Respond appropriately to an emergency (e.g., a broken leg). 10) Act according to my responsibilities to others (e.g., group members, parents/guardians, employers). 11) Look after the mountain environment and encourage others to do so too. 11) Look after the mountain environment and encourage others to do so too	11	11
*Number of QMDs logged by ideal self before assessment	SSHES	In an ideal world, how many Quality Mountain Days (QMDs) would you like to have logged before being assessed? This might be the minimum number (40) or it might be higher.	1	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Number of QMDs logged by ought self before assessment	SSHES	In an ideal world, how many Quality Mountain Days (QMDs) would you like to have logged before being assessed? This might be the minimum number (40) or it might be higher.	1	1
Efficacy to gain QMDs before being assessed	SSHES	To what extent do you agree with the following statement: "I know exactly what I need to do in order to gain all of the QMDs I would like to before being assessed."	1	1
Perception of assessment course staffs' "observation" skills	MCBS	The staff on my assessment course… 1) Paid close attention to what I did. 2) Carefully observed my skills.	4	2
Perception of assessment course staffs' "effective questioning" skills	MCBS	The staff on my assessment course… 1) Encouraged me to question the way I did things. 2) Encouraged me to make suggestions on how I could improve my performance.	4	2
Perception of assessment course staffs' "goal setting" skills	MCBS	The staff on my assessment course… 1) Helped me identify targets for attaining my goals. 2) Helped me set long term goals.	4	2
Perception of assessment course staffs' "developmental feedback" skills	MCBS	The staff on my assessment course… 1) Made sur I understood what I needed to do to improve. 2) Gave me advice on how to improve my skills.	4	2

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
Perception of assessment course staffs' "motivational feedback" skills	MCBS	The staff on my assessment course… 1) Expressed appreciation when I performed well. 2) Told me when I did a particularly good job.	4	2
Perceived provision of autonomy by assessment course staff	PESS	During my assessment course, the staff provided me with choices and options.	5	1
Perceived provision of structure by assessment course staff	PESS	During my assessment course, the staff made it clear to me what I need to do to get results.	5	1
Perceived involvement of assessment course staff	PESS	During my assessment course, the staff made enough time for me even though they were busy.	5	1
*Perceived relationship conflict between candidates on the assessment course	ICSS	For these statements, please think about the interactions between candidates (including yourself) on your assessment course. How much personality conflict was evident on your course?	5	1
*Perceived process conflict between candidates on the assessment course	ICSS	For these statements, please think about the interactions between candidates (including yourself) on your assessment course. To what extent did you disagree about the way to do things on your course?	4	1

Table B.4: Survey variables. *(continued)*

Variable	Measure	Item wording	Full	Short
*Perceived relationship conflict between staff and candidates on the assessment course	ICSS	Now we would like you to think about the interactions between candidates and the staff on your assessment course. How much personality conflict was evident on your course?	5	1
*Perceived process conflict between staff and candidates on the assessment course	ICSS	Now we would like you to think about the interactions between candidates and the staff on your assessment course. To what extent did you disagree about the way to do things on your course?	4	1
*Understanding of the standard required at assessment	SSHES	Some people talk about "the standard" required of candidates in order for them to become Mountain Leaders. For some people, the assessment course changes their view of the standard required to become a Mountain Leader, whereas for others, it confirms what they previously believed the standard to be. Following your assessment course, where did you feel the standard was in comparison to your perception of the standard prior to your assessment course? The standard was _____ I thought it was before the assessment course	1	1

* Included in final survey tool.

Appendix C

Chapter 3 Initial Classification Rates

Table C.1: Group 5 male candidates getting to assessment within 18 months of training, initial classification, training model performance.

Dataset	n	NB	SMO	IBk	J48	Rating
Psychosocial 2s	21	76.36	81.82	83.64	85.45	Very Good
Psychosocial 3s	11	76.36	80.00	78.18	85.45	Good
Training 2s	19	76.36	65.45	60.00	70.91	Modest
Training 3s	9	78.18	60.00	67.27	81.82	Good
Consolidation 2s	21	74.55	78.18	78.18	76.36	Good
Consolidation 3s	10	78.18	76.36	76.36	63.64	Good
DLOG t 2s	16	54.55	67.27	54.55	52.73	Poor
DLOG t 3s	7	56.36	65.45	49.09	60.00	Poor
DLOG t6 2s	17	61.82	56.73	50.91	58.18	Poor
DLOG t6 3s	7	65.45	56.55	63.64	63.64	Modest
DLOG t12 2s	17	61.82	56.73	50.91	58.18	Poor
DLOG t12 3s	7	65.45	56.55	63.64	63.64	Modest
DLOG t18 2s	17	74.55	67.27	70.91	67.27	Modest
DLOG t18 3s	7	76.36	72.73	70.91	67.27	Good
Merged survey 2s 2s	17	90.91	92.73	80.00	89.09	Very Good
Merged survey 2s 3s	11	87.27	89.09	80.00	90.91	Very Good
Merged survey 3s 2s	22	87.27	89.09	85.45	89.09	Very Good
Merged survey 3s 3s	16	85.45	92.73	85.45	89.09	Very Good
Merged DLOG 2s 2s	16	69.09	70.91	74.55	69.09	Modest
Merged DLOG 2s 3s	9	74.55	69.27	65.45	70.91	Good
Merged DLOG 3s 2s	20	69.09	70.00	69.09	69.09	Modest

Table C.1: Group 5 male candidates getting to assessment within 18 months of training, initial classification, training model performance. *(continued)*

Dataset	n	NB	SMO	IBk	J48	Rating
Merged DLOG 3s 3s	14	72.73	66.91	63.64	70.91	Modest
Merged 2s 2s	20	85.45	87.27	87.27	81.82	Very Good
Merged 2s 3s	10	89.09	92.73	90.91	87.27	Very Good
Merged 3s 2s	18	89.09	90.91	83.64	81.82	Very Good
Merged 3s 3s	11	87.27	87.45	87.27	85.45	Very Good
Centralised 2s	20	85.45	87.27	85.45	83.64	Very Good
Centralised 3s	7	85.45	89.09	90.91	89.09	Very Good

Table C.2: Group 5 female candidates getting to assessment within 18 months of training, initial classification, training model performance.

Dataset	n	NB	SMO	IBk	J48	Rating
Psychosocial 2s	20	85.19	87.04	79.63	88.89	Very Good
Psychosocial 3s	15	87.04	85.19	75.93	88.89	Very Good
Training 2s	20	72.22	66.67	66.67	66.67	Modest
Training 3s	10	68.52	68.33	66.67	64.81	Modest
Consolidation 2s	20	81.48	87.04	79.63	87.04	Very Good
Consolidation 3s	14	83.33	90.74	77.78	83.33	Very Good
Merged survey 2s 2s	21	84.07	86.30	87.31	80.28	Very Good
Merged survey 2s 3s	10	85.46	83.70	83.70	78.33	Very Good
Merged survey 3s 2s	23	87.50	92.86	85.71	78.57	Very Good
Merged survey 3s 3s	14	87.31	80.19	82.31	78.52	Very Good
DLOG t 2s	18	64.26	73.80	59.17	73.43	Modest
DLOG t 3s	13	67.69	66.67	62.87	71.67	Modest
DLOG t6 2s	13	66.39	75.19	71.85	58.70	Modest
DLOG t6 3s	16	73.52	72.22	70.19	54.81	Good
DLOG t12 2s	19	66.39	75.19	71.85	58.70	Modest
DLOG t12 3s	16	73.52	72.22	70.19	54.81	Good
DLOG t18 2s	20	66.67	75.74	66.67	66.67	Modest
DLOG t18 3s	13	75.93	73.52	66.67	61.11	Good
Merged DLOG 2s 2s	21	83.70	91.30	82.50	87.31	Very Good
Merged DLOG 2s 3s	12	87.13	87.69	84.35	91.02	Very Good
Merged DLOG 3s 2s	24	76.85	89.07	86.11	87.22	Very Good

Table C.2: Group 5 female candidates getting to assessment within 18 months of training, initial classification, training model performance. *(continued)*

Dataset	n	NB	SMO	IBk	J48	Rating
Merged DLOG 3s 3s	10	85.46	89.17	87.96	90.83	Very Good
Merged 2s 2s	22	94.54	92.69	90.93	83.52	Excellent
Merged 2s 3s	15	90.93	92.78	92.69	94.54	Excellent
Merged 3s 2s	22	92.69	94.44	88.98	87.50	Excellent
Merged 3s 3s	14	92.86	92.86	92.86	94.64	Excellent
Centralised 2s	19	90.93	92.31	94.54	87.50	Excellent
Centralised 3s	14	90.93	90.93	92.69	79.91	Excellent

Table C.3: Group 5 passing first time, initial classification, training model performance. Data standardised within sex.

Dataset	n	NB	SMO	IBk	J48	Rating
Psychosocial 2s	18	63.04	56.74	65.22	67.39	Modest
Psychosocial 3s	9	60.87	58.04	54.35	36.96	Poor
Training 2s	18	52.17	54.35	58.70	41.30	Poor
Training 3s	9	56.52	56.96	47.83	30.43	Poor
Consolidation 2s	19	60.87	65.43	60.87	41.30	Modest
Consolidation 3s	10	63.04	65.22	63.04	56.52	Modest
Merged survey 2s 2s	19	65.22	60.87	71.74	47.83	Modest
Merged survey 2s 3s	8	73.91	67.39	82.61	60.87	Good
Merged survey 3s 2s	23	52.17	66.74	58.70	54.35	Poor
Merged survey 3s 3s	10	50.00	73.26	60.87	39.13	Poor
DLOG t 2s	16	63.04	56.09	56.52	41.30	Poor
DLOG t 3s	9	56.52	61.96	52.17	50.00	Poor
DLOG a 2s	18	65.22	60.87	56.52	63.04	Modest
DLOG a 3s	5	58.70	29.57	69.57	63.04	Modest
Previous courses 2s	8	58.70	63.48	58.70	58.70	Poor
Previous courses 3s	6	56.52	57.83	52.17	30.43	Poor
Merged DLOG 2s 2s	23	60.87	52.39	63.04	65.22	Modest
Merged DLOG 2s 2s	23	63.04	62.83	71.74	73.91	Modest
Merged DLOG 3s 2s	12	56.52	59.13	56.52	56.52	Poor
Merged DLOG 3s 3s	7	63.04	49.13	63.04	63.04	Modest
Merged 2s 2s	22	65.22	69.35	71.74	47.83	Modest
Merged 2s 2s	22	69.57	60.87	65.22	50.00	Modest
Merged 3s 2s	24	60.87	48.04	58.70	54.35	Poor
Merged 3s 3s	11	60.87	57.39	60.87	54.35	Poor
Centralised 2s	18	76.09	69.57	82.61	60.87	Good
Centralised 3s	5	54.35	45.65	67.39	60.87	Poor

Appendix D

Expectations and Intentions

Results of Chapter 2 suggested that a significant number of candidates registered for the Mountain Leader qualification with the intention of only attending a training course and not going on to be assessed. However, this variable was not selected as an important discriminatory variable in the survey tool pilot work and was therefore dropped from the final survey tool. Given the incongruence of this finding with the results of Chapter 2, we inspected the data collected in the pilot work to better understand this discrepancy.

Candidates in Group 3 and Group 4 of the pilot data ($n = 532$) were asked, "Candidates who have registered for the ML may have different aspirations. Below is a list of common aspirations. Please tick the option which best reflected your aspirations at registration, you may only choose one." The response options were: (a) Mountain Leader training only, (b) Becoming a Mountain Leader, (c) Going onto higher walking qualifications, and (d) Going onto higher mountaineering qualifications. Interestingly, just 29 candidates (5.45%) from Groups 3 and 4 selected "Mountain Leader Training only." Further, nine of these candidates had been assessed 18 months after their training course. We would suggest that the rarity of the intention to only attend a training course in the pilot data is the reason it was not carried forward to the final survey tool. Furthermore, this rarity challenges the assumption of the participants in Chapter 2 as very few of the respondents to the pilot survey stated that they only intended to attend a training course and nearly a third of those who did state that had been assessed 18 months after their training course.

When considering this and the "context" findings of Chapter 3,[1] it appears that

[1]Which suggest the relative importance of becoming a Mountain Leader and, for male candidates, the strength of intention to be assessed at the end of a training course are important variables from discrim-

Table D.1: Summary statistics for candidates who had not been assessed when completing the survey.

Assessed	n	Female (%)	$M_{Age} \pm 1SD$	$M_{years\,since\,training} \pm 1$ SD	$M_{intention\,now} \pm 1SD$
FALSE	287	39.37	38.68 ± 12.34	0.87 ± 0.54	81.98 ± 27.65
TRUE	45	40.00	38.44 ± 11.78	0.75 ± 0.48	96.93 ± 9.96

the strength of intention is more important that the intention itself and we would hypothesise that the greater the strength of intention to be assessed, the greater the likelihood of being assessed. As well as asking candidates in the final survey group (i.e., those trained in 2017 and 2018) about the strength of their of intention to be assessed at registration, the start of training, and the end of training, we asked candidates who had not been assessed when they completed the survey ($n = 332$) to rate the strength of their intention to be assessed at that point in time on a scale from *no intention of being assessed* (0) to *every intention of being assessed* (100). This variable was not included in any of the pattern recognition analyses (see Table D.1 for summary statistics).

We performed a Bayesian t-test using the `BayesFactor` package (Morey and Rouder, 2018) in `R` (R Core Team, 2020), using the default settings, to test if there was a difference in the mean strength of intention to be assessed when completing the survey between those who had been assessed six months post-survey and those who had not been. Results of this prospective analysis showed strong evidence for their being a difference in the mean intention of being assessed between groups, $\mathrm{BF}_{10} = 62.37$. Figure D.1 shows the distribution of scores for both those who had been assessed and those who had not been. It is important to note that only three of the 45 candidates who had been assessed rated their intention lower than 90.

The analyses presented in this appendix suggest that most candidates do intend to be assessed but this intention must be strong for them to get to assessment. A potential implication of this finding could be that if Mountain Training wish to increase the number of candidates getting to assessment, course staff should aim to increase the strength of candidates' intentions to be assessed. Future studies could examine the efficacy and mechanism of any such intervention.

inating candidates who are assessed 18 months after their training course from those who are not.

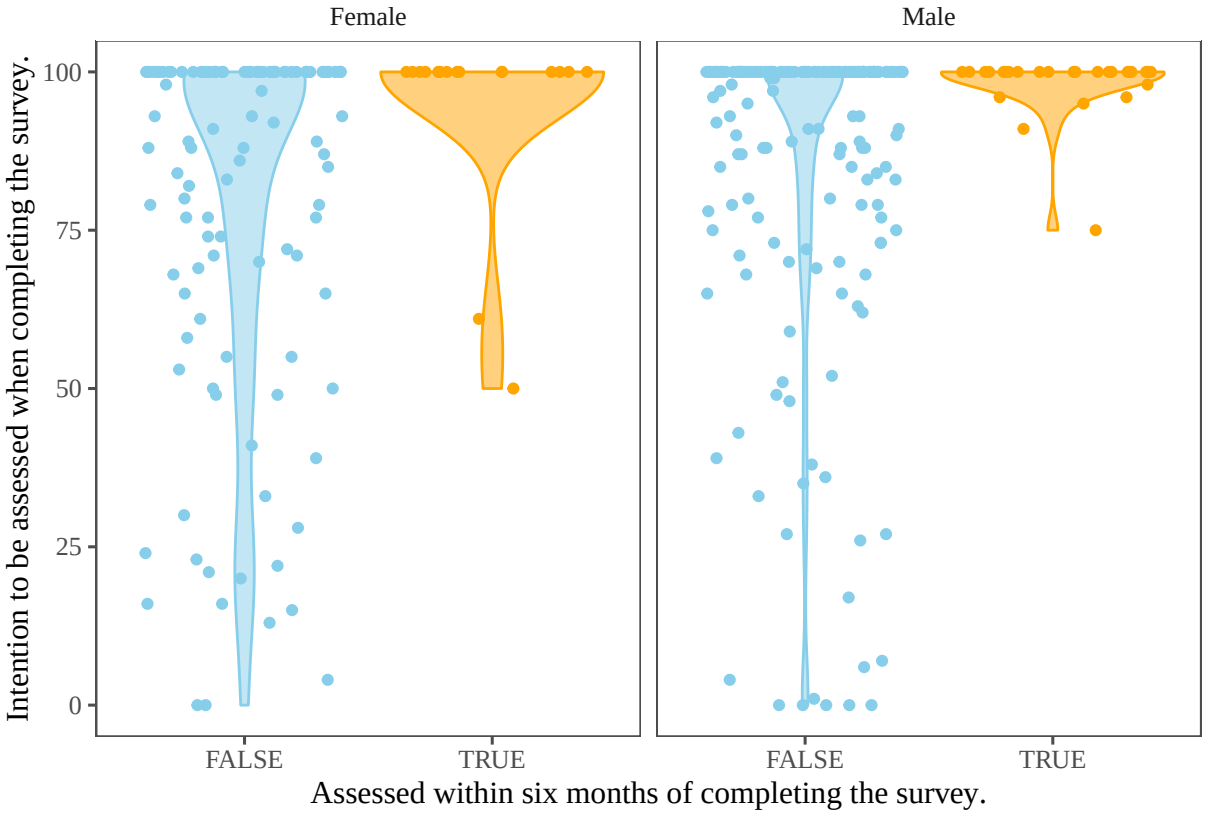

Figure D.1: Distribution of intention to be assessed when completing the survey by outcome with individual data points overlaid, grouped by sex.

Appendix E

Other Works Completed During the PhD

Basic and applied research with a variety of organisations

- Mountain Training organisations
  - MIA Survey 2016?
  - MCI study 2019 - designed study and collected data
  - MTS reanalysis of data
  - MTE impact survey and report
  - Financial impact of COVID-19
    - Report: https://www.mountain-training.org/latest-news/impact-of-covid-19-on-the-professional-mountaineering-community
    - Code: https://github.com/w-hardy/prof-mtnrs-covid-19
    - Montecarlo simulations
- UK Sport
  - Athlete perceptions of culture and success
  - Analysis of Situational Analysis and Sporting and Organisational Health documents
- England and Wales Cricket Board
- NRS Healthcare
- Public Health Wales

  - Glendinning et al. (2020)

- Innsbruck collaboration
  - (Frühauf et al., 2017, 2018)

- Dictionary entries
  - (Hardy and Woodman, 2019a,b)

- Robustness and resilience poster presentation at SPSP (Hardy et al., 2019)
  - Postgrad collaboration
  - https://osf.io/q4z3w/

- Statistics teaching
  - BSEM

References

Abernethy, B. (2013). Research informed practice. In Farrow, D., Baker, J., and MacMahon, C., editors, *Developing Sport Expertise: Researchers and Coaches put Theory into Practice*, pages 249–254. Routledge, London, 2nd edition.

Aha, D. W., Kibler, D., and Albert, M. K. (1991). Instance-based learning algorithms. *Machine Learning*, 6(1):37–66.

Aiken, L. S. and West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Sage, Newbury Park.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2):179–211.

Ajzen, I. and Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Prentice-Hall, Englewood Cliffs.

Ajzen, I. and Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, 22(5):453–474.

Allen, M. S., Greenlees, I., and Jones, M. (2013). Personality in sport: A comprehensive review. *International Review of Sport and Exercise Psychology*, 6(1):184–208.

American Psychological Association (2012). Guidelines for psychological practice with lesbian, gay, and bisexual clients. *American Psychologist*, 67(1):10–42.

American Psychological Association (2020). *Publication manual of the American Psychological Association.* American Psychological Association, Washington, DC, 7th edition.

Armitage, C. J. and Conner, M. (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. *British Journal of Social Psychology*, 40(4):471–499.

Baker, J. and Cobley, S. (2013). Outliers, talent codes and myths: Play and practice in developing the expert athlete. In Farrow, D., Baker, J., and MacMahon, C., editors, *Developing Sport Expertise: Researchers and Coaches put Theory into Practice*, pages 13–25. Routledge, London, 2nd edition.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2):191–215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37(2):122–147.

Bandura, A. (1986). *Social foundations of thought and action*. Prentice-Hall, Englewood Cliffs.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman, New York.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In Pajares, F. and Urdan, T. C., editors, *Self-efficacy beliefs of adolescents*, chapter 14, pages 307–337. Information Age Publishing, Greenwich.

Bandura, A. and Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology*, 45(5):1017–1028.

Beattie, S., Hardy, L., Savage, J., Woodman, T., and Callow, N. (2011). Development and validation of a trait measure of robustness of self-confidence. *Psychology of Sport and Exercise*, 12(2):184–191.

Beattie, S., Hardy, L., and Woodman, T. (2015). A longitudinal examination of the interactive effects of goal importance and self-efficacy upon multiple life goal progress. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 47(3):201–206.

Bell, J. J., Hardy, L., and Beattie, S. (2013). Enhancing mental toughness and performance under pressure in elite young cricketers: A 2-year longitudinal intervention. *Sport, Exercise, and Performance Psychology*, 2(4):281–297.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2):238–246.

Biddle, S., Markland, D., Gilbourne, D., Chatzisarantis, N. L. D., and Sparkes, A. C. (2001). Research methods in sport and exercise psychology: quantitative and qualitative issues. *Journal of sports sciences*, 19:777–809.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Bolón-Canedo, V., Sánchez-Maroño, N., and Alonso-Betanzos, A. (2015a). Distributed feature selection: An application to microarray data classification. *Applied Soft Computing Journal*, 30:136–150.

Bolón-Canedo, V., Sánchez-Maroño, N., and Alonso-Betanzos, A. (2015b). Recent advances and emerging challenges of feature selection in the context of big data. *Knowledge-Based Systems*, 86:33–45.

Bouckaert, R. R., Frank, E., Hall, M. A., Kirkby, R., Reutemann, P., Seewald, A., and Scuse, D. (2018). *WEKA Manual for Version 3-9-3*. University of Waikato, Hamilton.

Boulter, M., Hardy, J., Roberts, R., and Woodman, T. (3001). Intra-group conflict scale for sport.

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

Brown, I. and Inouye, D. K. (1978). Learned helplessness through modeling: The role of perceived similarity in competence. *Journal of Personality and Social Psychology*, 36(8):900–908.

Bueno, J., Weinberg, R. S., Fernández-Castro, J., and Capdevila, L. (2008). Emotional and motivational mechanisms mediating the influence of goal setting on endurance athletes' performance. *Psychology of Sport and Exercise*, 9(6):786–799.

Caraway, K., Tucker, C. M., Reinke, W. M., and Hall, C. (2003). Self-efficacy, goal orientation, and fear of failure as predictors of school engagement in high school students. *Psychology in the Schools*, 40(4):417–427.

Cassidy, S. and Eachus, P. (2002). Developing the computer user self-efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computing Research*, 26(2):133–153.

Cervone, D. and Palmer, B. W. (1990). Anchoring biases and the perseverance of self-efficacy beliefs. *Cognitive Therapy and Research*, 14(4):401–416.

Chantal, Y., Guay, F., Dobreva-Martinova, T., and Vallerand, R. J. (1996). Motivation and elite performance: An exploratory investigation with Bulgarian athletes.

Chase, M. A. (2001). Children's self-efficacy, motivational intentions, and attributions in physical education and sport. *Research Quarterly for Exercise and Sport*, 72(1):47–54.

Chemolli, E. and Gagné, M. (2014). Evidence against the continuum structure underlying motivation measures derived from self-determination theory. *Psychological Assessment*, 26(2):575–585.

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., and Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, 39(2):216–236.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3):464–504.

Cho, J. Y. and Lee, E.-H. (2014). Reducing confusion about Grounded Theory and qualitative content analysis: Similarities and differences. *The Qualitative Report*, 19(32):1–20.

Coppock, A. (2019). *randomizr: Easy-to-Use Tools for Common Forms of Random Assignment and Sampling*. R package version 0.20.0.

D'Alpinisme, U. I. d. A. (2015). UIAA Standards for voluntary leaders and instructors.

Dawson, K. A. and Brawley, L. R. (2000). Examining the relationship between exercise goals, self-efficacy, and overt behavior with beginning exercisers. *Journal of Applied Social Psychology*, 30(2):315–329.

de Rooij, M. and Weeda, W. (2020). Cross-Validation: A Method Every Psychologist Should Know. *Advances in Methods and Practices in Psychological Science*, page 251524591989846.

de Wit, F. R. C., Greer, L. L., and Jehn, K. A. (2012). The paradox of intragroup conflict: A meta-analysis. *Journal of Applied Psychology*, 97(2):360–390.

Deci, E. L. (1975). *Intrinsic Motivation*. Plenum Press, Boston, MA.

Deci, E. L. and Ryan, R. M. (1980). Self-determination theory: When mind mediates behavior. *The Journal of Mind and Behavior*, 1(1):33–43.

Deci, E. L. and Ryan, R. M. (1985a). *Intrinsic motivation and self-determination in human behavior*. Plenum, New York.

Deci, E. L. and Ryan, R. M. (1985b). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2):109–134.

Deci, E. L. and Ryan, R. M. (1991). A motivational approach to self: integration in personality. In Dienstbier, R., editor, *Nebraska Symposium on Motivation. Perspectives on Motivation*, volume 38, pages 237–288. University of Nebraska Press, Lincoln, NE.

Deci, E. L. and Ryan, R. m. (1995). Human autonomy: The basis for true self-esteem. In Kernis, M. H., editor, *Efficacy, Agency, and Self-Esteem*, pages 31–49. Springer, Boston, MA.

Deci, E. L. and Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4):227–268.

Deci, E. L. and Ryan, R. M. (2014). Autonomy and Need Satisfaction in Close Relationships: Relationships Motivation Theory. In *Human Motivation and Interpersonal Relationships*, volume 9789401785, pages 53–73. Springer Netherlands, Dordrecht.

Denzin, N. K. and Lincoln, Y. S. (2005). *The SAGE Handbook of Qualitative Research*. SAGE Publications Ltd., London, 3rd edition.

Duda, R. O., Hart, P. E., and Stork, D. G. (2000). *Pattern Classification*. John Wiley & Sons, 2nd edition.

Evan, W. M. (1965). Conflict and performance in R&D organizations. *Industrial Management Review*, 7(1):37–46.

Fishbein, M. (2000). The role of theory in HIV prevention. *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV*, 12(3):273–278.

Fornell, C. and Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1):39–50.

Fortier, M., Vallerand, R., Brière, N., and Provencher, P. (1995). Competitive and recreational sport structures and gender: A test of their relationship with sport motivation. *International Journal of Sport Psychology*, 26:24–39.

Frank, E., Hall, M. A., and Witten, I. H. (2016). *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*. Morgan Kaufmann, 4th edition.

Freeman, P., Coffee, P., Moll, T., Rees, T., and Sammy, N. (2014). The ARSQ: The Athletes' Received Support Questionnaire. *Journal of Sport and Exercise Psychology*, 36(2):189–202.

Freeman, P., Coffee, P., and Rees, T. (2011). The PASS-Q: The Perceived Available Support in Sport Questionnaire. *Journal of Sport and Exercise Psychology*, 33(1):54–74.

Freeman, P. and Rees, T. (2009). How does perceived support lead to better performance? An examination of potential mechanisms. *Journal of Applied Sport Psychology*, 21(4):429–441.

Frost, R. O., Marten, P., Lahart, C., and Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research*, 14(5):449–468.

Frühauf, A., Hardy, W. A. S., Pfoestl, D., Hoellen, F.-G. G., and Kopp, M. (2017). A Qualitative Approach on Motives and Aspects of Risks in Freeriding. *Frontiers in Psychology*, 8(November):1–13.

Frühauf, A., Hardy, W. A. S., Roberts, R., Niedermeier, M., and Kopp, M. (2018). Structural validation of three German versions of behavioral and motivational scales in high-risk sports. *German Journal of Exercise and Sport Research*, 48(4):467–477.

Gagné, F. (2004). Transforming gifts into talents: the DMGT as a developmental theory. *High Ability Studies*, 15(2):119–147.

Glendinning, F., Barton, E. R., Newbury, A., McManus, M. A., Janssen, H., Johnson, G., Ramos Rodriguez, G., Harker, S., and Bellis, M. A. (2020). An evaluation of the Adverse Childhood Experience Trauma Informed Multi-agency Early Action Together (ACE TIME) training: national roll out to police and partners. Technical report, Public Health Wales.

Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. In Hewstone, M. and Strobee, W., editors, *European review of social psychology*, pages 141–185. Wiley, Chichester.

Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6):504–528.

Gould, D. (2005). Goal setting for peak performance. In Williams, J., editor, *Applied sport psychology: Personal growth to peak performance*. Mayfield, Palo Alto, CA, 5th edition.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, 6(4):430–450.

Güllich, A., Hardy, L., Kuncheva, L. I., Laing, S., Evans, L., Rees, T., Abernethy, B., Côté, J., Warr, C., and Wraith, L. (2019). Developmental biographies of Olympic Super-Elite and Elite Athletes: A multidisciplinary Pattern Recognition Analysis. *Journal of Expertise*, 2(1):23–46.

Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182.

Guyon, I., Weston, J., and Barnhill, S. (2002). Gene selection for cancer classification using Support Vector Machines. *Machine Learning*, 46:389–422.

Hagger, M. S. and Chatzisarantis, N. L. D. (2009). Integrating the theory of planned behaviour and self-determination theory in health behaviour: A meta-analysis. *British Journal of Health Psychology*, 14(2):275–302.

Hagger, M. S. and Chatzisarantis, N. L. D. (2014). An integrated behavior change model for physical activity. *Exercise and Sport Sciences Reviews*, 42(2):62–69.

Hagger, M. S. and Chatzisarantis, N. L. D. (2015). Self-Determination theory. In Conner, M. and Norman, P., editors, *Predicting and changing health behaviour: Research and practice with social cognition models*, number January, chapter 4, pages 107–141. Open University Press, Maidenhead, 3rd edition.

Hagger, M. S., Chatzisarantis, N. L. D., and Biddle, S. (2002). A Meta-analytic review of the theories of reasoned action and planned behaviour in physical activity: Predictive validity and the contribution of additional variables. *Journal of Sport & Exercise Psychology*, 24:3–32.

Hamilton, K., Kirkpatrick, A., Rebar, A., and Hagger, M. S. (2017). Child sun safety: Application of an integrated behavior change model. *Health Psychology*, 36(9):916–926.

Hardy, L., Barlow, M., Evans, L., Rees, T., Woodman, T., and Warr, C. (2017). Great British medalists: Psychosocial biographies of Super-Elite and Elite athletes from Olympic sports. *Progress in Brain Research*, 232:1–119.

Hardy, L., Jones, G., and Gould, D. (1996). *Understanding psychological preparation for sport*. John Wiley & Sons, Chichester.

Hardy, W. A. S., Glendinning, F., Barwood, C., and Boulter, M. (2019). The Robustness and Resilience Questionnaire: Initial validation [Poster presentation]. Society of Personality and Social Psychology Conference.

Hardy, W. A. S. and Woodman, T. (2019a). Risk-taking. In Hackfort, D., Schinke, R. J., and Strauss, B., editors, *Dictionary of Sport Psychology*, chapter R, pages 251–252. Elsevier, London.

Hardy, W. A. S. and Woodman, T. (2019b). Sensation Seeking. In Hackfort, D., Schinke, R. J., and Strauss, B., editors, *Dictionary of Sport Psychology*, chapter S, page 268. Elsevier.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning. *Bayesian Forecasting and Dynamic Models*, 1:1–694.

Hauck, L. J., Carpenter, M. G., and Frank, J. S. (2008). Task-specific measures of balance efficacy, anxiety, and stability and their relationship to clinical balance performance. *Gait and Posture*, 27(4):676–682.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94(3):319–340.

Hopwood, M. J. (2013). *The Developmental History of Athletes Questionnaire: Towards a comprehensive understanding of the development of sport expertise*. Doctoral, Victoria University.

Horvath, S. and Röthlin, P. (2018). How to Improve Athletes' Return of Investment: Shortening Questionnaires in the Applied Sport Psychology Setting. *Journal of Applied Sport Psychology*, 30(2):241–248.

Hu, L.-t. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis : Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1):1–51.

Hutchinson, J. C., Sherman, T., Martinovic, N., and Tenenbaum, G. (2008). The effect of manipulated self-efficacy on perceived and sustained effort. *Journal of Applied Sport Psychology*, 20(4):457–472.

Ingledew, D. K., Markland, D., and Ferguson, E. (2009). Three Levels of Exercise Motivation. *Applied Psychology: Health and Well-Being*, 1(3):336–355.

Ingledew, D. K., Wray, J. L., Markland, D., and Hardy, L. (2005). Work-related goal perceptions and affective well-being. *Journal of Health Psychology*, 10(1):101–122.

Jehn, K. A. (1995). A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. *Administrative Science Quarterly*, 40(2):256.

John, G. H. and Langley, P. (1995). Estimating Continuous Distributions in Bayesian Classifiers. In *Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 338–345, San Mateo.

Johnston, K., Wattie, N., Schorer, J., and Baker, J. (2018). Talent Identification in Sport: A Systematic Review. *Sports Medicine*, 48(1):97–109.

Jones, B. D. (2019). *'Game Changers': Discriminating Features within the Microstructure of Practice and Developmental Histories of Super-Elite Cricketers - A Pattern Recognition Approach.* Doctoral, Bangor University.

Jones, B. D., Hardy, L., Lawrence, G. P., Kuncheva, L. I., Brandon, R., Bobat, M., and Thorpe, G. (2020). It ain't what you do-It's the way that you do it: Is optimizing challenge key in the development of super-elite batsmen? *Journal of Expertise*, 3(2):144–168.

Jones, B. D., Hardy, L., Lawrence, G. P., Kuncheva, L. I., Du Preez, T. L., Brandon, R., Such, P., and Bobat, M. (2019). The identification of 'Game Changers' in England cricket's developmental pathway for elite spin bowling: a machine learning approach. *Journal of Expertise*, 2(2):92–120.

Jones, G. (1995). More than just a game: Research developments and issues in competitive anxiety in sport. *British Journal of Psychology*, 86(4):449–478.

Jones, G. and Swain, A. (1992). Intensity and direction as dimensions of competitive state anxiety and relationships with competitiveness. *Perceptual and Motor Skills*, 74(2):467–472.

Jöreskog, K. G. (1989). *LISREL 7: User's reference guide.* Scientific, Mooresvill, IN.

Judge, T. A., Bono, J. E., Ilies, R., and Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4):765–780.

Kasser, T. and Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin*, 22(3):280–287.

Kernan, M. C. and Lord, R. G. (1990). Effects of valence, expectancies, and goal-performance discrepancies in single and multiple goal environments. *Journal of Applied Psychology*, 75(2):194–203.

Kessler, R. C. (2003). Scoring the HPQ.

Kessler, R. C., Barber, C., Beck, A., Berglund, P., Cleary, P. D., McKenas, D., Pronk, N., Simon, G., Stang, P., Ustun, T. B., and Wang, P. (2003). The World Health Organization Health and Work Performance Questionnaire (HPQ). *Journal of Occupational and Environmental Medicine*, 45(2):156–174.

Kieslich, P. J., Wulff, D. U., Henninger, F., and Haslbeck, J. M. B. (2019). *mousetrap: Process and Analyze Mouse-Tracking Data*. R package version 3.1.4.

Kira, K. and Rendell, L. A. (1992). Feature selection problem: traditional methods and a new algorithm. *Proceedings Tenth National Conference on Artificial Intelligence*, pages 129–134.

Klassen, R. M. and Chiu, M. M. (2010). Effects on Teachers' Self-Efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress. *Journal of Educational Psychology*, 102(3):741–756.

Kline, R. B. (2016). *Principles and practice of structural equation modeling.* Guilford Press, New York, 4th edition.

Kuncheva, L. I. and Rodríguez, J. J. (2018). On feature selection protocols for very low-sample-size data. *Pattern Recognition*, 81:660–673.

Kyllo, L. B. and Landers, D. M. (1995). Goal Setting in Sport and Exercise: A Research Synthesis to Resolve the Controversy. *Journal of Sport and Exercise Psychology*, 17(2):117–137.

Levy, P. (1968). Short-form tests: A methodological review. *Psychological Bulletin*, 69(6):410–416.

Lincoln, Y. S. and Guba, E. G. (1985). *Naturalistic inquiry.* Sage, Beverly Hills, CA.

Lirgg, C. D. and Feltz, D. L. (1991). Teacher versus peer models revisited: Effects on motor performance and self-efficacy. *Research Quarterly for Exercise and Sport*, 62(2):217–224.

Little, B. R. (1983). Personal projects: A rationale and method for investigation. *Environment and Behavior*, 15(3):273–309.

Locke, E. A., Frederick, E., Lee, C., and Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology*, 69(2):241–251.

Locke, E. A. and Latham, G. P. (1990). *A theory of goal setting and task performance*. Prentice-Hall, Englewood Cliffs.

Maddux, J. E. (1995). Self-efficacy theory: An introduction. In Maddux, J. E., editor, *Self-efficacy, adaptation, and adjustment: Theory, research, and application.*, pages 3–33. Plenum Press, New York.

Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences*, 42(5):851–858.

Markland, D. and Tobin, V. J. (2010). Need support and behavioural regulations for exercise among exercise referral scheme clients: The mediating role of psychological need satisfaction. *Psychology of Sport and Exercise*, 11(2):91–99.

Martin, J. J. (2002). Training and performance self-efficacy, affect, and performance in wheelchair road racers. *Sport Psychologist*, 16(4):384–395.

McAuley, E., Konopack, J. F., Motl, R. W., Morris, K. S., Doerksen, S. E., and Rosengren, K. R. (2006). Physical activity and quality of life in older adults: Influence of health status and self-efficacy. *Annals of Behavioral Medicine*, 31(1):99–103.

McCrae, R. R. and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1):81–90.

McCullagh, P. and Weiss, M. R. (2001). Modeling: considerations for motor skill performance and psychological responses. In Singer, R. . N., Hasenblaus, H. A., and Janelle, C. M., editors, *Handbook of sport psychology*, pages 205–38. Wiley, New York, 2nd edition.

Meade, A. W. and Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4):361–388.

Means, B. and Loftus, E. F. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology*, 5(4):297–318.

Miller, M. A. and Rahe, R. H. (1997). Life changes scaling for the 1990s. *Journal of Psychosomatic Research*, 43(3):279–292.

Miller, R. L. and Mulligan, R. D. (2002). Terror management: the effects of mortality salience and locus of control on risk-taking behaviors. *Personality and Individual Differences*, 33:1203–1214.

Morey, R. D. and Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for Common Designs*. R package version 0.9.12-4.2.

Moritz, S. E., Feltz, D. L., Fahrbach, K. R., and Mack, D. E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport*, 71(3):280–294.

Mountain Training (2019). Mountain Leader FAQs.

Mountain Training England (2019). Impact report 2018. Technical report.

Mountain Training UK (2015a). Mountain Leader - Candidate Handbook.

Mountain Training UK (2015b). Mountain Leader Award: Skills checklist.

Mountain Training UK (2019). Mountain Training: Candidate Management System.

Multon, K. D., Brown, S. D., and Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation.

Murphy, C. A., Coover, D., and Owen, S. V. (1989). Development and Validation of the Computer Self-Efficacy Scale. *Educational and Psychological Measurement*, 49(4):893–899.

Muthén, B. and Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3):313–335.

Niemiec, C. P., Ryan, R. M., and Deci, E. L. (2009). The path taken: Consequences of attaining intrinsic and extrinsic aspirations in post-college life. *Journal of Research in Personality*, 43(3):291–306.

Orbell, S., Johnston, M., Rowley, D., Davey, P., and Espley, A. (2001). Self-efficacy and goal importance in the prediction of physical disability in people following hospitalization: A prospective study. *British Journal of Health Psychology*, 6(1):25–40.

Patton, M. Q. (2002). *Qualitative research and evaluation methods.* SAGE Publications, Thousand Oaks, CA, 3rd edition.

Pearson, D. T., Naughton, G. A., and Torode, M. (2006). Predictability of physiological testing and the role of maturation in talent identification for adolescent team sports. *Journal of Science and Medicine in Sport*, 9(4):277–287.

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., and Brière, N. M. (2001). Associations among perceived autonomy support, forms of self-regulation, and persistence: A prospective study. *Motivation and Emotion*, 25(4):279–306.

Platt, J. C. (1998). Fast training of Support Vector Machines using Sequential Minimal Optimization. In Schoelkopf, B., Burges, C., and Smola, A., editors, *Advances in Kernel Methods - Support Vector Learning*. MIT Press, Cambridge.

Putnick, D. L. and Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41:71–90.

QSR (2017). Nvivo 11 Pro.

Qualtrics (2019). Qualtrics.

Quinlan, R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rees, T. and Freeman, P. (2007). The effects of perceived and received support on self-confidence. *Journal of Sports Sciences*, 25(9):1057–1065.

Rees, T., Hardy, L., Güllich, A., Abernethy, B., Côté, J., Woodman, T., Montgomery, H., Laing, S., and Warr, C. (2016). The Great British Medallists Project: A Review of Current Knowledge on the Development of the World's Best Sporting Talent. *Sports Medicine*, 46:1041–1058.

Rees, T., Ingledew, D. K., and Hardy, L. (1999). Social support dimensions and components of performance in tennis. *Journal of Sports Sciences*, 17(5):421–429.

Reunanen, J. (2003). Overfitting in making comparisons between variable selection methods. *Journal of Machine Learning Research*, 3:1371–1382.

Roberts, R., Woodman, T., Hardy, L., Davis, L., and Wallace, H. M. (2013). Psychological Skills Do Not Always Help Performance: The Moderating Role of Narcissism. *Journal of Applied Sport Psychology*, 25(3):316–325.

Rosseel, Y., Jorgensen, T. D., and Rockwood, N. (2020). *lavaan: Latent Variable Analysis*. R package version 0.6-7.

Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of Personality*, 63(3):397–427.

Ryan, R. M. and Connell, J. P. (1989). Perceived Locus of Causality and Internalization: Examining Reasons for Acting in Two Domains. *Journal of Personality and Social Psychology*, 57(5):749–761.

Ryan, R. M., Connell, J. P., and Deci, E. L. (1985). A motivational analysis of self-determination and self-regulation in education. In Ames, C. and Ames, R. E., editors, *Research on motivation in education: The classroom milieu*, pages 13–51. Academic Press, New York.

Ryan, R. M. and Deci, E. L. (2000a). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1):54–67.

Ryan, R. M. and Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78.

Ryan, R. M. and Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In Deci, E. L. and Ryan, R. M., editors, *Handbook of self-determination research*, pages 3–33. University of Rochester Press, Rochester.

Ryan, R. M. and Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press, New York.

Ryan, R. M. and Deci, E. L. (2019). Brick by Brick: The Origins, Development, and Future of Self-Determination Theory. In Elliot, A. J., editor, *Advances in Motivation Science*, volume 6, pages 111–156. Elsevier Inc., Cambridge, MA.

Ryan, R. M., Deci, E. L., and Vansteenkiste, M. (2016). Autonomy and autonomy disturbances in self-development and psychopathology: research on motivation, attachment, and clinical process. In Cicchetti, D., editor, *Developmental Psychopathology*, pages 385–438. John Wiley & Sons, Hoboken, 3rd edition.

Samson, A. and Solmon, M. (2011). Examining the sources of self-efficacy for physical activity within the sport and exercise domains. *International Review of Sport and Exercise Psychology*, 4(1):70–89.

Satorra, A. and Bentler, P. (2001). is used as a chi-square test statistic with degrees of freedom equal to the difference on the number of independent parameters estimated under the models AdO and Adl. As in the case of the goodness-of-fit test, it is of interest to scale the statistic. *Psychometrika*, 66(4):507–514.

Schmidt, A. M. and DeShon, R. P. (2007). What to Do? The Effects of Discrepancies, Incentives, and Time on Dynamic Goal Prioritization. *Journal of Applied Psychology*, 92(4):928–941.

Schmidt, A. M. and Dolis, C. M. (2009). Something's Got to Give: The Effects of Dual-Goal Difficulty, Goal Progress, and Expectancies on Resource Allocation. *Journal of Applied Psychology*, 94(3):678–691.

Schmidt, A. M., Dolis, C. M., and Tolli, A. P. (2009). A Matter of Time: Individual Differences, Contextual Dynamics, and Goal Progress Effects on Multiple-Goal Self-Regulation. *Journal of Applied Psychology*, 94(3):692–709.

Schorer, J. and Elferink-Gemser, M. (2013). How good are we at predicting athletes' futures? In Farrow, D., Baker, J., and MacMahon, C., editors, *Developing Sport Expertise: Researchers and Coaches put Theory into Practice*, pages 30–39. Routledge, London, 2nd edition.

Seery, M. D. and Quinton, W. J. (2016). Understanding Resilience: From Negative Life Events to Everyday Stressors. In *Advances in Experimental Social Psychology*. Elsevier.

Sheldon, K. M. and Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, 76(3):482–497.

Sheldon, K. M. and Kasser, T. (1998). Pursuing Personal Goals: Skills Enable Progress, but Not all Progress is Beneficial. *Personality and Social Psychology Bulletin*, 24(12):1319–1331.

Sheldon, K. M. and Prentice, M. (2019). Self-determination theory as a foundation for personality researchers. *Journal of Personality*, 87(1):5–14.

Sitzmann, T. and Yeo, G. (2013). A meta-analytic investigation of the within-person self-efficacy domain: Is self-efficacy a product of past performance or a driver of future performance? *Personnel Psychology*, 66(3):531–568.

Smialowski, P., Frishman, D., and Kramer, S. (2010). Pitfalls of supervised feature selection. *Bioinformatics*, 26(3):440–443.

Smith, B. and McGannon, K. R. (2018). Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. *International Review of Sport and Exercise Psychology*, 11(1):101–121.

Smith, B. and Sparkes, A. C. (2016). Introduction: An invitation to qualitative research. In Taylo, editor, *Handbook of Qualitative Research in Sport and Exercise*, number September. Taylor and Francis.

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., and Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back. *International journal of behavioral medicine*, 15(3):194–200.

Smith, G. T., McCarthy, D. M., and Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1):102–111.

Sniehotta, F. F., Presseau, J., and Araújo-Soares, V. (2014). Time to retire the theory of planned behaviour. *Health Psychology Review*, 8(1):1–7.

Snowdonia National Park Authority (2017). Mountain rescue incidents on Snowdon: An analysis of hill walking incidents 2009-2016. Technical report.

Sparkes, A. C. (1998). Validity in Qualitative Inquiry and the Problem of Criteria: Implications for Sport Psychology. *The Sport Psychologist*, 12(4):363–386.

Sparkes, A. C. and Smith, B. (2009). Judging the quality of qualitative inquiry: Criteriology and relativism in action. *Psychology of Sport and Exercise*, 10(5):491–497.

Sparkes, A. C. and Smith, B. (2014). *Qualitative Research Methods in Sport, Exercise and Health: From Process to Product*. Routledge, Oxon.

Sport England (2020). Active Lives Adult Survey November 2018/19 Report. Technical Report April.

Stajkovic, A. D. and Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124(2):240–261.

Steiger, J. H. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*, 25(2):173–180.

Tabachnick, B. and Fidell, L. (2007). *Using Multivariate Statistics*. Allyn and Bacon, New York, 5th edition.

Tenenbaum, G., Hall, H. K., Calcagnini, N., Lange, R., Freeman, G., and Lloyd, M. (2001). Coping with physical exertion and negative feedback under competitive and self-standard conditions. *Journal of Applied Social Psychology*, 31(8):1582–1626.

Theeboom, T., Beersma, B., and van Vianen, A. E. M. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*, 9(1):1–18.

Torkzadeh, G. and Koufteros, X. (1994). Factorial validity of a computer self-efficacy scale and the impact of computer training. *Educational and Psychological Measurement*, 54(3):813–821.

Vallerand, R. J. (1997). Toward A Hierarchical Model of Intrinsic and Extrinsic Motivation. In Zanna, M. P., editor, *Advances in Experimental Social Psychology*, volume 29, pages 271–360. Academic Press, San Diego, CA, US.

Vallerand, R. J. and Blssonnette, R. (1992). Intrinsic, Extrinsic, and Amotivational Styles as Predictors of Behavior: A Prospective Study. *Journal of Personality*, 60(3):599–620.

Vallerand, R. J. and Ratelle, C. F. (2002). Intrinsic and extrinsic motivation: A hierarchical model. In *Handbook of self-determination research*, pages 37–63. University of Rochester Press, Rochester.

van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., and Muthén, B. (2013a). Facing off with Scylla and Charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology*, 4:1–15.

van de Schoot, R., Yerkes, M. A., Mouw, J. M., and Sonneveld, H. (2013b). What Took Them So Long? Explaining PhD Delays among Doctoral Candidates. *PLoS ONE*, 8(7).

Visa, S., Ramsay, B., Ralescu, A., and van der Knapp, E. (2011). Confusion Matrix-based Feature Selection. In *23rd MAI & CS Conference*, Cincinnati.

Wagstaff, C. R. D., Arthur, C. A., and Hardy, L. (2018). The Development and Initial Validation of a Measure of Coaching Behaviors in a Sample of Army Recruits. *Journal of Applied Sport Psychology*, 30(3):341–357.

Walker, C. O., Greene, B. A., and Mansell, R. A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences*, 16(1):1–12.

Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence*. Psychological Corporation, New York.

Weinberg, R., Gould, D., and Jackson, A. (1979). Expectations and Performance: An Empirical Test of Bandura's Self-efficacy Theory. *Journal of Sport Psychology*, 1(4):320–331.

Weinberg, R. S. and Gould, D. (2014). *Foundations of sport and exercise psychology*. Human Kinetics, Champaign, IL, 6th edition.

Weinberg, R. S., Gould, D., Yukelson, D., and Jackson, A. (1981). The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task. *Journal of Sport Psychology*, 3(4):345–354.

Weinberg, R. S., Yukelson, D., and Jackson, A. (1980). Effect of Public and Private Efficacy Expectations on Competitive Performance. *Journal of Sport Psychology*, 2(4):340–349.

Yu, L. and Liu, H. (2003). Feature selection for high-dimensional data: a fast correlation-based filter solution. In *Proceedings of the twentieth international conference on machine learning*, pages 865–863.

Yuan, K.-H. and Bentler, P. M. (1997). Mean and Covariance Structure Analysis: Theoretical and Practical Improvements. *Journal of the American Statistical Association*, 92(438):767.

Yukl, G., Kim, H., and Chavez, C. (1999). Task importance, feasibility, and agent influence behavior as determinants of target commitment. *Journal of Applied Psychology*, 84(1):137–143.

Yukl, G., Kim, H., and Falbe, C. M. (1996). Antecedents of influence outcomes. *Journal of Applied Psychology*, 81(3):309–317.