Cite this paper as:

Neill, J. T., Marsh, H. W., & Richards, G. E. (in preparation). The Life Effectiveness Questionnaire: Development and psychometrics. Sydney: University of Western Sydney.

Chapter 7:
The Life Effectiveness Questionnaire:
Development and Psychometrics (Study 1)

Draft 7x

Word count: ~10,000

28 June, 2001

7.1 Abstract 2
7.2 Introduction 3
7.2.1 Rationale for the Life Effectiveness Questionnaire (LEQ) 3
Length and Complexity 4
Relevance to Program Aims 5
Assessment of Competence 5
Sensitivity to Change 5
Educational Exercise 5
7.2.2 Stages of Development 6
7.2.3 Dimensions of Life Effectiveness 7
Time Management 7
Social Competence 7
Achievement Motivation 8
Intellectual Flexibility 8
Organisational Self-Discipline 8
Productive Teamwork 8
Task Leadership 9
Emotional Control 9
Active Initiative 9
Self Confidence 9
Hardiness Resourcefulness 10
Overall Structure 10
7.3 Method 12
Procedure 12
Statistical Analyses 12
Congeneric Analysis 12
Multifactorial Analysis 12
Invariance testing 13
Outcome Statistics 15
7.4 Results and Discussion 16
Study 1 16
Congeneric Analyses 16
The Initial Baseline Model 17
Model Development 17
Invariance Testing 19
Study 2 19
Congeneric Analyses 19
Multi-factorial CFA and Invariance Testing 20
Invariance Testing between Study 1 and Study 2 21
Summary and Conclusion 22

7.1 Overview

This chapter briefly reviews the problematic history of evaluating the outcomes of outdoor-based experiential education programs. One of the key criticisms is the lack of theoretically and psychometrically sound instruments to assess changes in the self-perceptions of participants in such programs. Consequently, this chapter explains the development of a practical, easy-to-use self-report instrument designed to measure changes in key areas of "life effectiveness". Congeneric and multifactorial confirmatory factor analyses were applied to the Life Effectiveness Questionnaire-version G to develop an eight factor model (Time Management, Social Competence, Achievement Motivation, Intellectual Flexibility, Task Leadership, Emotional Control, Active Initiative, and Self-Confidence). Factorial invariance across gender and age groups was supported. A replication study confirmed the structure of the instrument. Potential applications of the Life Effectiveness Questionnaire for program evaluations are also discussed.

7.2 Introduction

"Let man be happy, informed, skillful, well-behaved, and productive."
- B. F. Skinner (1953)

The present investigation has pragmatic, theoretical, and methodological purposes. The pragmatic goal is to produce a short, easy-to-administer, self-report "life effectiveness" instrument with sound psychometric properties. The theoretical emphasis is to identify and explore psychological and behavioural domains which constitute "life fitness" or "life proficiency" and which are theoretically amenable to developmental change though experience-based educational interventions. The methodological focus is to demonstrate a process of instrument development through the application of confirmatory factor analysis (CFA), including congeneric, multifactorial, and invariance analyses.

Firstly, a rationale for the Life Effectiveness Questionnaire (LEQ) is presented, including a brief explanation of each of the hypothesized domains. Secondly, an explanation of CFA methodology is provided. Finally, the results of the instrument are presented and discussed, with suggested directions for further research using the Life Effectiveness Questionnaire (LEQ).

7.2.1 Rationale for the Life Effectiveness Questionnaire (LEQ)

Experiential and outdoor education practitioners have been making official claims that challenging outdoor expeditions are beneficial for participant's personal development at least since organisations such as Outward Bound began in Wales in 1942. For many years such claims were considered by enthusiasts and proponents of outdoor education to be vindicated by their personal experience, anecdotal successes, and a notion that challenging experience-based activities, usually in the outdoors, were inherently beneficial. Skeptics could be understandably concerned that such claims were over zealous because, until the 1960's, only a few outcome studies with weak methodologies had been conducted to evaluate the outcomes of such programs.

The period from the early 1960's to the 1980's saw more rigorous evaluation of outdoor experience-based program outcomes. An important strength of these studies was the use of psychometrically developed instrumentation such as Rotter's Locus Of Control (Rotter, 1966) (e.g. Collingwood, 1972; Gaston, 1978; Marsh & Richards, 1985; Wright, 1982), the Coopersmith Self-Esteem Inventory (Coopersmith, 1984) (e.g. Porter, 1975; Richards & Richards, 1981), the Tennessee Self-Concept Scale (Fitts, 1965) (e.g. Marsh & Richards, 1988a; Nye, 1975; Wetmore, 1972; Wright, 1982), and the 16 Personality Factors (Cattell, Eber, & Tatsuoka, 1970) (e.g. Hendy, 1975; Owens & Richards, 1979).

This era of studies produced many interesting insights into the effects of adventure education programs and provided a useful foundation for ongoing research (Ewert, 1982); Richards, 1977). A limitation of these studies was that the instruments were not specifically designed to measure changes, were at times only indirectly matched to the program aims, and did not always have psychometric structures which stood

the test of applied research settings (e.g. see review by Marsh, Richards, & Barnes, 1986a,b).

In the 1980's and 1990's, a valuable series of studies evaluated the effects of Outward Bound programs on multidimensional self-concept using Marsh's Self-Description Questionnaire I for children (Marsh & Richards, 1988b), SDQ II for adolescents (Neill, 1994) and SDQ III for young adults (Marsh, Richards & Barnes, 1986a,b; Mitchell & Mitchell, 1989). The use of psychometrically sound multi-dimensional instruments which could be linked more closely to the aims of outdoor and experience-based program, included larger sample sizes to give greater power, and more rigorous methodologies, such as truncated time series designs, were positive advancements made by these studies.

Notwithstanding these research efforts, it is interesting that published research on adventure education program outcomes has overwhelmingly used instruments primarily designed for assessment purposes and not necessarily designed for measuring developmental changes. This has largely been the case because there is no readily available and rigorously developed instruments specifically designed for measuring the types of personal changes likely to occur as a result of adventure education programs. The absence of such instrumentation is somewhat surprising, particularly given the expansion of outdoor education into school curriculums, management training, and the proliferation of experiential and outdoor-based education organisations across the world.

Formal evaluation of these programs is becoming a more common practice today than in decades past, yet the field is still plagued by its traditional appeal to anecdotal evaluations and the use of instrumentation inadequately matched to program aims or insensitive to measuring change. These, among other methodological issues such as low statistical power (e.g. see Hattie, Marsh, Neill, & Richards, 1997), could well be the causes of frustration expressed by outdoor education practitioners with inconsistent results from empirical research projects. Practitioners are often personally convinced that their programs are achieving valuable benefits for participants, which may well be the case, but the methodological problems of off-the-shelf instrumentation studies may produce an empirical non-result. Clients, stakeholders, and policy makers are in need of appropriate instrumentation and accompanying methodology to provide the means for better evaluation of program outcomes and to make comparisons between the relative efficacy of different educational approaches.

The present study describes the development of a self-report measure of key areas of life effectiveness that are likely to be targeted by personal development intervention programs. Although developed within the context of adventure programs, the instrument is intended for broader use with experiential education and personal development intervention programs. Given this overall objective, there were five principles which guided the instrument's development.

Length and Complexity

On outdoor and experience-based programs, instruments are often administered in field settings (e.g. in the bush, on board a boat, in various weather conditions), on multiple occasions (e.g. pre-program, first day of program, last day of program and

post-program followup) and to a wide range of participants (e.g. people with learning disabilities, people without English as their first language, school children, corporate managers). Hence, the shorter and simpler an instrument (reliability and validity aside), the greater the instrument's potential applicability. The aim was to develop an instrument which would provide a maximum amount and type of unique information in as short a time as possible (i.e. a maximum of about ten minutes). The instrument's instructions and layout also needed to be straight-forward to allow people without research experience, such as group leaders or teachers, to administer the instrument consistently across different groups.

Relevance to Program Aims

Generally, a major aim of many outdoor experiential programs is to facilitate individuals' personal development in a broad range of life skills (e.g. self-confidence, initiative, communication skills, etc.), although different programs may have more specific aims such as the development of teamwork and leadership skills. Ideally, the instrument would encompass a wide range of life proficiency domains relevant to general and specific program aims, so as to allow for within and between program comparison of different program outcomes.

Assessment of Competence

The range of life skills assessed by the instrument should be necessary, or at least beneficial, to effective and successful living and working. The greater the specificity of a measure, the more likely it is to predict actual performance (Blau, 1993). Hence, the aim was for the instrument items to focus on self-assessment of competence at practical skills which are relevant to a broad range of personal and professional situations.

Sensitivity to Change

The scoring system is an important aspect of an instrument's sensitivity to change. A dichotomous (yes/no) scoring does not provide much sensitivity, whereas a large range can reduce the instrument's reliability. A balance needs to be reached between sensitivity to change and reliability. Despite being a critical issue, a search of the literature revealed little research, for example, on the relative efficacy of likert-type scales with different numbers of responses for measuring change.

Two further issues to be considered in developing an instrument for measuring change are ceiling/floor effects and test-retest correlations. It is desirable that the wording of the items and the response scale tends to produce responses from participants that leave room for detecting shifts in self-perceptions either up or down -- hence the means need to be examined during the instrument development. In addition, test-retest correlations give an indication of the stability of items and scales. If these correlations are low, then participants' responses to the item or scale may change for reasons other than those that can be attributed to an intervention experience.

Educational Exercise

The methods used to facilitate personal change during outdoor experiential programs include providing opportunities for self-assessment, goal-setting, and feedback on

personal progress. An instrument which can be used to facilitate the processes of self-examination, goal-setting and feedback would give it added value.

7.2.2 Stages of Development

The concept of "personal effectiveness" rarely appears in psychological and educational literature and is poorly researched. "Life skills" appears, however it usually refers to the basic functional skills required by remedial students or people with disabilities to function as independent members of society.
Williams, Eyring, Gaynor, and Long (1991) developed a *Views of Life Scale* which measures self-perceptions on three scales, including one called "effective life management", described as follows:

> One's life is well-managed - One believes s/he is doing those things that make the best use of his or her resources. Sh/e believes that s/he has opportunities for self-improvement, focuses on the present, makes good decisions, solves problems efficiently, achieves desired goals, maintains a balance among the important areas of life, does mainly what s/he enjoys, and is managing life with increased effectiveness (p.168).

An oblique factor analysis of their initial data, however, revealed five factors, two of which had considerable overlap with the "effective life management scale" - Quality of Life (e.g. "I have lots of/little energy") and Life Management (e.g. "I usually/seldom achieve my most important goals"). The *Views of Life Scale* was developed for health and well-being assessment. The structure of the "effective life management scale" as a unitary factor was not strong and the two factor model clearly needed much work and refinement.

This study took a different approach and considered "life effectiveness" (or "life proficiencies" or "life skills") as the psychological and behavioural aspects of human functioning which determine a person's effectiveness or proficiency in any given situation. On closer examination of the literature, a number of potential life effectiveness dimensions were identified. Most of these potential dimensions had been studied in relative isolation from the other potential dimensions. Hence, much work needed to be done to examine how the potential life effectiveness dimensions were interrelated.

Five major stages of pilot testing preceded the present study (which covers Stages 6 and 7; see Table 1). For each successive stage:

a) Separate exploratory factor analyses, where sample sizes were sufficient, were conducted for male and female participants to help ensure that a gender-biased structure was not developed.
b) Generally speaking, larger sample sizes were used as the study progressed (N = *105* for Stage 1 through to N = 1164 for Stage 6 and N = 960 for Stage 7).
c) The background of participants in successive stages tended to be increasingly diverse, to help ensure wide applicability for the instrument.

Insert Table 1 about here

7.2.3 Dimensions of Life Effectiveness

Following the first five stages of pilot instruments, the LEQ-G was developed, which had eight factors from previous versions and three new factors. These eleven dimensions are described below and summarized in Table 2.

Time Management

Increasingly in Western society, an individual's ability to plan and make efficient use of time is seen as a useful quality in both personal and professional life. A multitude of popular psychology articles and books (e.g., Covey & Merrill, 1994; Taylor & Mackenzie, 1986) attest to this growing interest in time management and how it can be improved. Despite this growing awareness and expanding lay literature, surprisingly little empirical research on time management has been conducted (Macan, 1994). Studies that have been conducted tend to be focused on the work-place (e.g., Buehler, Griffin, & Ross, 1994; Hall & Hursch, 1982; Macan, 1994) or applied to narrowly defined populations (e.g. King, Winett, & Lovett, 1986). There is a distinct need to develop instrumentation to assess the time management proficiency of people in everyday life and for the effectiveness of time management intervention strategies to be evaluated.

Social Competence

* Social skills, it can also be argued are a core component of being generally effective in personal and professional life. Much psychological theory also attests to the extent to which social skills are learnt via modeling and that they can be enhanced through positive intervention.
# Interpersonal intelligence (Gardner, 19*)
Social competence is also, theoretically, trainable
cell – social skills training programs
#"Small group participation in enjoyable outdoor adventure activities appears to improve individual socialisation…Shared participation in challenging adventure activities can result in improved understanding, trust and communication between young participants and between young people and adult staff."
- Barret & Greenaway (1995a, p. 51).

This dimension seeks to encompass the ability of an individual to function effectively in social situations. In the literature this is also referred to as Interpersonal Competence and Social Skills (*references – see saved search file). Empirical studies have established social competence as having a factor structure distinguishable from other aspects of personality and behaviour (Pelechano, de Miguel, & Penate, 1991). Different aspects of social competence have been examined in the research literature (e.g. peer social competence (Inderbitzen, 1994)) and the social competence for different populations has been investigated (e.g. young children (Saunders & Green, 1993) and the elderly (Pelechano, et al., 1991)). In the LEQ, Social Competence items applicable to a wide range of settings and a diverse range of participants (e.g., school children through to corporate executives) were sought.

Achievement Motivation

The research literature has established a strong link between achievement and motivation (Arkes, 1982). Basically, the more a person is motivated to achieve, the more likely it is that the person will reach a particular achievement (Weiner, 1980). Consequently, intervention efforts have been able to improve achievement levels by enhancing motivation to achieve (Nicholls, 1984).

Traditional measurements of Achievement Motivation have been projective, using techniques such as the Thematic Apperception Test (TAT; McClelland, *). In the TAT, participants express their interpretations of picture cards, which provide skilled researchers with information about the participants' motivations. The major limitation of the TAT is that it is time-consuming to administer and hence is impractical for large-scale administration. Other measures such as sentence completion tests have attempted to simplify the projective procedure (French, 1955). In addition, more cognitive and direct measures of Achievement Motivation have been developed. McClelland (1985) criticizes these approaches for being overly cognitive and failing to take into account the physiological processes underlying motivation which, he claims, are more reliably accessed through the TAT. Despite these reservations of McClelland's, practicality led to a cognitive-based self-report format for the LEQ Achievement Motivation scale.

Intellectual Flexibility

Intellectual Flexibility refers to the ability of a person to appropriately adjust their views to accommodate and act upon the ideas of others. Intellectual Flexibility scales have appeared in other instruments such as the Personality Research Form (Jackson, 1984; the scale is called Cognitive Structure). Links between Intellectual Flexibility have been established with existential openness, that is, the tendency to confront existential issues such as certainty-uncertainty, purposefulness-meaninglessness, and self-determination-fate (Stevens, 1992) with conceptual complexity (Suedfeld, 1995).

Organisational Self-Discipline

The Organisational Self-Discipline dimension refers to the effectiveness of a person's planning, organisation, and subsequent action to ensure that tasks are well organised and completed. Although there is no substantive research evidence available to support the existence of Organisational Self-Discipline as a distinct factor, the notion of self-discipline and self-organisation appears in many self-help books and personal development courses, hence common experience suggests *prima facie* validity for such as scale . On the strength of its presence in lay literature related to life effectiveness, an Organisational Self-Discipline scale was created for the LEQ.

Productive Teamwork

Personal and working lives regularly involve individuals pursuing tasks with other people. In these situations, a key element of an individual's effectiveness is his/her ability to be productive through collaboration. According to Dyer (1984) past attempts to measure teamwork have been problematic. Dyer contended that many measures of teamwork are unreliable, complex, insensitive, and measure irrelevant variables. The LEQ Productive Teamwork scale does not attempt to measure a

team's functioning, but instead aims to assess an individual's own evaluation of his/her ability in a range of interpersonal values and behaviours which are likely to enhance his/her effectiveness as a team member.

Task Leadership

The majority of research literature on leadership focuses on the characteristics of people performing designated leadership roles, such as managers in organisations (e.g. Mumford, 1986), or on situational factors leading to the emergence of leaders in experimental laboratory studies (e.g. Zaccaro, Foti, & Kenny, 1991). In contrast, the LEQ Task Leadership scale gets individuals to assess their ability to take on and perform in a leadership role when there is a situational need or opportunity. An individual who is able to take control of situations, motivate and enthuse others towards common goals, and ensure a productive and harmonious outcome, is more likely to be effective in general life than a person unable to perform such functions.

Emotional Control

Emotion has a long history of being assessed in philosophical and psychological literature (Goleman, 1995). Measures of emotional competence, are included in personality instruments (e.g. High School Personality Questionnaire (Cattell, 1968)), clinical instruments (e.g. Beck Depression Inventory (Beck, Steer &, Garbin, 1988)), and coping instruments (see Stanton, Danoff-Burg, Cameron, & Ellis (1994) for a review). There appears to be good evidence that emotional control is associated with effective life outcomes. For example, Kanner and Feldman (1991) demonstrated that control over both emotional hassles and particularly emotional uplifts in early adolescence is associated with resistance to depression and greater behavioural restraint. Too much emotional control, which may function as a psychological defensive mechanism, however could be detrimental. The focus for the LEQ Emotional Control scale was for participants to assess their ability to deal with emotions under difficult or demanding situations. This focus was partly based on theoretical grounds - that such situations are salient tests of a person's emotional control - and partly for the practical reason that outdoor programs often aim to positively effect people's personal coping in difficult situations through the controlled and progressive application of challenging tasks and situations based in the outdoors.

Active Initiative

Active Initiative is intended to capture the dynamic ability that is demonstrated by an individual who actively and independently initiates new actions and thoughts in a variety of personal and work settings. It follows that Active Initiative should be relevant to both productive and dynamic performance at work and to taking control of and directing one's personal life. Despite this prima facie validity based on common experience, there is currently little research literature available to test the idea that Active Initiative can be considered a component of life effectiveness. This was a new dimension proposed for the LEQ version G.

Self Confidence

Self Confidence in a term used in everyday language to refer to an individual's general belief in his/her abilities. In contrast, the research literature tends not to use the general term self confidence, instead focusing more specifically on areas such as

self-esteem, self-efficacy, self-concept, and so on. It is not clear which of these areas most closely represents the popular notion of Self Confidence. For the purposes of the LEQ, it was hypothesized that there could be an identifiable general personal confidence component of personal effectiveness. The Self Confidence scale was intended to provide a self-assessed measure of an individual's general confidence of success in work and personal situations. This was a new dimension proposed for the LEQ version G.

Hardiness Resourcefulness

Hardiness describes a collection of personal attributes which combine to allow a person to function effectively in difficult or demanding situations. Kobasa (Antonovsky, 1987; Kobasa, 1979; Maddi & Khoshaba, 1994) sees hardiness as composed of the interrelated self-perceptions of commitment, control, and challenge. These characteristics are helpful in allowing a person to experience potentially stressful circumstances in developmental rather than debilitative ways.

Accumulating evidence suggests that Hardiness is a reasonable measure of mental health and general well-being (Maddi & Khoshaba, 1994), a buffer between stresses and illnesses (Antonovsky, 1987; Kobasa, Maddi, & Kahn, 1982), associated with adaptive coping behaviours (Williams, Wiebe, & Smith, 1992), moderately associated with optimism (Scheier & Carver, 1985; Sweetman, Munz, & Wheeler, 1993), and appears to increase effectiveness in situations such as sporting performance (Maddi & Hess, 1992) and success in training situations (Westman, 1990).

A second characteristic, 'resourcefulness', has received less attention but appears to be related to 'hardiness'. Resourcefulness is the ability to maintain productivity and effectiveness with less-than-ideal personal, social, and environmental resources available. Rosenbaum (1990) has shown that resourcefulness is linked to adaptively coping with various stressors.

For the LEQ, a Hardiness Resourcefulness scale is proposed which incorporates the two areas into a unitary dimension measuring individual's perception of their ability to be resourceful and hardy, particularly when faced with difficult situations.

Overall Structure

Finally, a hypothesized structure for the LEQ instrument was considered. One of the major aims of the LEQ is to identify, evaluate, and determine change in key aspects of life effectiveness, hence the factors should be structurally distinct. It is expected that the factors will, however, be moderately intercorrelated, since they are all hypothesized to pertain to the notion of 'life effectiveness'. However, highly correlated factors do not provide unique information and consequently strong consideration should be given to dropping redundant items and factors, particularly given the desire for a short instrument. Alternatively, if the hypothesized factors are not distinct, then a single, global 'life effectiveness' factor could provide a reasonable model fit. It is planned to test a single factor model to help decide whether the proposed multidimensional structure is appropriate. A global measure of

life effectiveness would have practical value for comparative evaluation of the effectiveness of personal development intervention programs.

Given the intended use of the LEQ for evaluating educational programs outcomes, it is desirable that its structure is the same for males and females and the same across age groups. Marsh (1993b) reports that factorial invariance is a much neglected aspect of self-concept instrument development. Hence a contribution of this study is to demonstrate invariance testing methodology.

7.3 Method

Procedure

Administration of the LEQ-G (Study 1) and the subsequent version, the LEQ-H (Study 2), took place in a wide variety of field settings, mostly on the first day of various outdoor experience-based programs, as well as some single administrations to available groups. The instruments were administered by program coordinators or group leaders using a common, written protocol which included asking participants to follow the written instructions on the questionnaire.

Statistical Analyses

Confirmatory Factor Analysis (CFA)

All statistical analyses were conducted with the commercially available 'PC' version of LISREL 7 (Jöreskög & Sörbom, 1989). Detailed explanation of the full Confirmatory Factor Analysis (CFA) procedure is available elsewhere (e.g. Bollen 1989; Jöreskög & Sörbom, 1989; Pedhazur & Schmelkin, 1991), however a discussion of the elements of CFA as applied to instrument development in the current study warrant discussion.

CFA evaluates the degree of fit between the variance-covariance matrix of a structural model of researcher-defined parameters and an observed variance-covariance matrix based on actual data. In the present study three types of CFA are used - congeneric analyses, multifactorial CFAs, and multifactorial CFAs to test for structural invariance between groups.

Congeneric Analysis

The first type of CFA is the congeneric analysis which are conducted separately for each proposed LEQ scale. Items from a scale, such as Time Management, are tested for fit to a single factor model. An important assumption underlying the construction of each of the LEQ scales is that the set of items designed to measure each scale is a unidimensional construct. The one-factor congeneric model provides a direct test of this assumption. In addition, congeneric measurement models allow for differences in the degree to which each individual item contributes to the overall composite scale score to be assessed, which is very useful for refining the items in the scale. The outcomes of the congeneric analyses are considered alongside outcomes for the multifactorial CFAs, providing the researcher with sound psychometric information which, when used in conjunction with substantive considerations, as the basis for making decisions about which items and dimensions to keep, modify, or drop.

Multifactorial Analysis

Following congeneric analyses, a multifactorial CFA of the instrument's initial baseline model is conducted. The multifactorial CFA tests the adequacy of the putative scales for describing the variance in the data. The modification indices (MIs) provided by LISREL (Jöreskög & Sörbom, 1989) indicate how much improvement in model fit would be obtained for each parameter if its loading was not specified and allowed to be free. When using CFA to develop an instrument the MIs of most interest are those for the item factor loadings, although the researcher should be also be cognizant of MIs which indicate that the unique error for two items are correlated because this may indicate a method effect such as for positive and

negative item wording. Also of interest for instrument development are the factor correlations. In this study, while expecting moderate correlations between the factors, these correlations should be low enough that each scale measures a substantially different area from each of the other scales. Factor correlations below .5 were considered desirable and above .7 were not acceptable.

Following examination of the outcomes for the congeneric analyses and the multifactorial initial baseline CFA, alternative models of the data are proposed and tested. Given the pragmatic aim to develop a psychometrically sound, quick-to-administer instrument, modifications to the initial baseline model are likely to involve dropping items which do not load cleanly on a putative factor, or dropping factors which do not have a distinct set of cleanly loading items. Decisions to drop items and factors should be informed by careful examination of the outcome statistics (see below for a detailed description of these statistics) and must be evaluated substantively - for example, "Can the instrument still reasonably fulfill its overall purpose without this item or factor?". The refined model of the data is then retested with another multifactorial CFA, and the process continues until a satisfactory model is established. Particularly when substantial modifications have been made, it is also important to then test the final model with new data.

Described as it is here, instrument development using multifactorial CFA may not appear dramatically different to instrument development using exploratory factor analysis (EFA). There are, however, some critical advantages of CFA over EFA. CFA's major advantage is that parameters such as factor loadings can be fixed or allowed to be free, according to the model specifications, providing that there is substantive justification for doing so, whereas in EFA all factor loadings are automatically free. In addition, CFA provides the researcher with valuable psychometric diagnostics such as the MIs, and other outcome statistics, particularly indices of the goodness-of-fit (described below) which provide overall measures of the degree of improvement in model fit for each successive model adjustment. Hence, CFA offers far greater power for fine tuning an instrument's development than is afforded by EFA.

Invariance testing

A major intended use of the Life Effectiveness Questionnaire for program evaluation is to compare mean differences between groups, such as between males and females, between adults and adolescents, or between participants participating in different intervention programs. Before such comparisons can be meaningfully made, the structural invariance of the instrument must be tested. An instrument is structurally invariant if different subpopulations respond to the instrument using a similar perception of the measurement scale (Drasgow & Kanfer, 1985). Structural invariance means that different subpopulations respond to an instrument using a similar implicit perception of the number of factors, the loadings of items on those factors, and similar variances and covariances amongst the items. Knowing that an instrument is structurally invariant allows the researcher or program evaluator to draw meaningful conclusions about mean differences between groups.

Before testing for structural invariance, it is important to consider *a priori* knowledge of group differences in the structure of the psychological constructs (Byrne, 1989). If there is reason to believe that invariance between groups may not

exist, then development of different structural models should be considered. In the case of the LEQ, there were no expectations for structural variance between males and females or between the age groups in this sample, and hence proceeded to develop a structural model intended to be the same for all groups.

One methodological strategy for testing structural equivalence between groups is the LISREL model already discussed above. The first step is to establish a well-fitting and theoretically viable model (Byrne, Shavelson, & Muthén, 1989; Reise, Widaman, & Pugh, 1993). This model is then used to examine similarities and differences in the covariance patterns across the subpopulations of interest (Windle, Iwawaki, & Lerner, 1988).

A global test of structural invariance evaluates whether or not all the structural model parameters are equivalent across groups. While seemingly straightforward, this global test can often lead to findings which are contradictory with subsequent tests for partial structural invariance (Byrne, Shavelson, & Muthén, 1989). This suggests that conducting tests of partial structural invariance in addition to the test of full structural invariance is prudent. Hence, the current study tests the hypotheses that there is invariance for:

1. the number of factors,
2. the factor-loading pattern (i.e., $\underline{H}_0$: $\Lambda_1 = \Lambda_2 = ... \Lambda_G$),
3. the factor variances and covariances (i.e., $\underline{H}_0$: $\Phi_1 = \Phi_2 = ... \Phi_G$),
4. the error/uniquenesses (i.e., $\underline{H}_0$: $\theta_1 = \theta_2 = ... \theta_G$).

The tenability of hypotheses 1 and 2 are logical prerequisites to hypotheses 3 and 4. Support for hypotheses 1 and 2 is termed factorial invariance and may be all that is substantively necessary. The invariance of factor variances and covariances (hypothesis 3) and the error/uniquenesses (hypothesis 4) is termed total invariance and it is of interest, but of less central concern since these parameters and their relationship across groups will typically be idiosyncratic to a particular sample (MacCallum & Tucker, 1991; Reise et al., 1993) unless the samples are taken randomly from the same population (Marsh, 1993a).

For invariance testing, the data is split into the requisite groups and the model is re-tested with the constraint that the parameters of the above hypotheses are held invariant across the groups. Because all multi-group models imposing invariance constraints over groups are nested under models with no such invariance constraints, the different in chi-squares for the two models can be evaluated for statistical significance in relation to the difference in degrees of freedom under appropriate conditions. However, this test of statistical significance is subject to all the limitations of other tests of statistical significance that have been largely abandoned in favor of fit indices in CFA research (Marsh, 1993). For this reason, the focus should be more on fit indices in evaluating support for the invariance of parameter estimates. The fit indices for this invariance model are then calculated. If these fit indices are close to the fit indices without invariance constraints then the researcher can be confident that there is invariance of the model being tested. If, on the other hand, there is a substantial worsening of the fit indices then this indicates that the parameters being tested are not invariant across the groups.

A further application of this technique, demonstrated in this chapter, is that it can be used to test for invariance between responses to different instruments which contain common items. It is possible, for example, that the contextual effects of the surrounding items effect the way in which participants respond to individual items (Marsh, 1997?*). If those same individual items are placed in the context of different surrounding items, different patterns of relationships amongst the items may result. A typical situation where this might occur is when a subset of items is taken from a larger instrument. Whether or not participants' responses to the relationships between the items are effected by the change in context can be examined by using invariance testing. A particular advantage of this approach, providing that invariance can be shown, is that data for common items from different instruments can be legitimately pooled for subsequent analyses.

Outcome Statistics

There are a number of outcome statistics produced in CFA which are useful tools for scale and instrument development:

a) Factor Loadings (λ_x).
The square of an item's factor loading (λ_x) in standardised form equals the percentage of its variance that is accounted for by the latent factor. As a general guide, it is desirable for items to have at least half of their variance explained by the latent factor. Hence, a factor loading of .7 (49% of variance explained) is a common criterion for acceptable items.

b) Co-efficient Omega (Ω).
The co-efficient omega is a measure of how well the item variables serve jointly as a measurement instrument of the latent construct. In other words, the co-efficient omega measures scale reliability. The co-efficient omega is a lower bound estimate of the true reliability to the extent that any of the constituent items have unique error. As a general guide, a co-efficient omega over .7 is minimally acceptable and over .8 is desirable (*ref needed?).

c) Fit Indices (FIs).
A number of outcome statistics can be calculated to measure the degree to which the model fits the data. The traditional outcome statistic is the chi-square (χ^2) which is smaller for better-fitting models. However, the χ^2 statistic is problematic in that it increases with the degrees of freedom (*df*) and the sample size (*n*) when the model fit is imperfect[1]. Also problematic for the χ^2 is that if the observed variables in a sample do not have a multivariate normal distribution or are not continuous variables, then the χ^2 may also be inflated (Drasgow & Kanfer, 1985). As a result of inflated χ^2 values, significant probabilities (*p*) are reported, and possibly good models can be rejected if this criterion is used as the fit index.

Due to the obvious problems with χ^2 and *p* for model evaluation many other indices of model fit have been developed, with much discussion of their relative merits (e.g., Bollen, 1990; Bentler, 1989, 1990; Cudeck & Browne, 1983; Jöreskög & Sörbom,

[1]In practice, a model fit is always imperfect, unless every parameter is freely estimated so that the model is 'perfectly overfitted' to the data.

1989; Marsh, Balla, & McDonald, 1988; McDonald & Marsh, 1990). This study adopts the approach argued for and demonstrated by Marsh (1993a, 1994). This involves the Tucker-Lewis index (TLI) (Tucker & Lewis, 1973) (also termed the non-normed fit index (NNFI) by Bentler & Bonnet, 1980) because it is relatively independent of sample size and was found by Marsh, Balla, and McDonald (1988) to be among the best of the then-available practical indices of fit. The TLI measures a model fit relative to the fit of a null[2] model, hence it will be unstable if applied to a data set well-described by a null model (McDonald & Marsh, 1990). The relative noncentrality index (RNI) is also reported. The RNI also indicates roughly the proportion of covariation among indicators explained by the CFA model relative to a null model.

For both the TLI and RNI, values near 0 indicate poor fit and values near 1 indicate good fit. Values greater than .90 are usually considered satisfactory, although values closer to .95 are desirable. The TLI and RNI differ in that the TLI has a correction for a lack of model parsimony, whereas the RNI does not. As a result it is possible for the imposition of invariance constraints to result in no decrement -- or even an increase -- in the TLI and this is taken as support for the invariance constraints (Marsh, 1993a,b). It is possible to calculate other fit indices such as the parsimony index based on the RNI from the values presented.

7.4 Results and Discussion

Study 1

Congeneric Analyses

The outcome statistics for the one-factor congeneric analyses on the eleven LEQ-G scales are presented in Table 3. All scales had co-efficient omegas above .8, with the lowest clearly being for Organisational-Self Discipline (Ω=.829). This scale was also the only one with average factor loadings below .7 (*?.696).

Six scales appear to fit single factor models very well as indicated by TLIs and RNIs over .950 (except for the Social Competence RNI of .924). These are the 7-item Time Management scale and the 5-item scales Social Competence, Achievement Motivation, Intellectual Flexibility, Task Leadership, and Emotional Control.

[2]A null model is a 'non-model' in the sense that it is does not specify any theoretical structure for the data. In other words, the null model does not test a structural model to account for the variances and covariances of the items, and thereby does not achieve any data reducation. The null model specifies that each item loads loads perfectly on a separate factor, so there are the same number of factors as there are items. The value of the null model is that it serves as a baseline for assessing the relative improvement in goodness-of-fit when other models when other models are applied to the data.

The three scales which were new to the LEQ-version G (Active Initiative, Self Confidence, and Hardiness Resourcefulness) are promising, with TLIs and RNIs over .90. Closer examination of the weaker items in the multidimensional CFA will help refine these scales.

Finally, two scales (Organisational Self-Discipline and Productive Teamwork) are problematic, with TLIs below .80. These scales require closer examination in the multidimensional CFA and may need to be dropped from the life effectiveness model.

The Initial Baseline Model

Turning to the multifactorial CFAs of the initial baseline 64-item, 11-factor a priori LEQ-G model. Listwise deletion produced a sample size of 1164 cases (672 males, 484 females) from an original sample size of 1405 cases. The parameter estimates of the CFA are presented in Table 4 and the factor variances and covariances are in Table 5.
The fit indices in Table 6 indicate that the initial baseline 11-factor model provides a much better model for the data than the null model, however the initial baseline model is rejected as an adequate model fit, with fit indices well under .9.

Model Development

The next step was to refine the initial baseline model. Four out of the eleven scales were considered developmental when designing the LEQ-G, and consequently had more than five items. Hence, the first step was to reduce each of these scales to five items by removing items which made the least contribution to defining their respective factor in the congeneric CFAs and which had relatively high modification indices indicating a lack of "purity" in the initial baseline multifactorial CFA. This new 55-item, 11-factor model provided a better fit to the data, however it was still unsatisfactory, with all the fit indices below .9 (see Table 6).
The next refinement involved examining problematic factors. The Organisational Self-Discipline scale, the existence of which had been hypothesized although it had not appeared previously in the research literature, did not perform well in the one-factor congeneric analyses and in the multifactorial CFAs produced high MIs, particularly for cross-loadings on the Time Management factor. This makes substantive sense - that one's management of time is closely related to one's organisation and self-discipline - borne out by a very high factor correlation between the two scales of .95 (see Table 5). With this amount of covariation there was clearly little to distinguish the two sets of items from each other. With the lowest co-efficient omega and lowest average factor loadings, it was obvious that the Organisational Self-Discipline items represented a less viable construct than Time Management, and so they were dropped. The resultant 50-item 10-factor model was tested and showed improvements in all the indices of fit to approximately .9 (see Table 6).

Further examination of the one-factor congeneric analyses and the multifactorial MIs indicated that most scales contained at least one more weak item. The worst fitting item, it terms of congeneric and multi-factorial CFA factor loadings and modification indices, was eliminated from each of the 10 remaining factors. This 40-item, 10-

factor model showed further improvements in the TLI and RNI, pushing them both over .9, but still not to the desired .95 level (see Table 6).

Further examination of the factor structure for the 40-item 10-factor model and the MIs showed that the Hardiness Resourcefulness scale was problematic. Although the congeneric analysis of the Hardiness Resourcefulness scale had appeared promising, in the context of the other scales the multifactorial CFA showed that the Hardiness Resourcefulness items tended to have high modification indices for factor loadings on the Emotional Control and Self Confidence scales and tended to have moderate modification indices for factor loadings on the Intellectual Flexibility and Productive Teamwork scales. This problem was also indicated in the factor correlations (in the initial baseline model) between Hardiness Resourcefulness and Self Confidence ($\underline{r}$ = .86) and Hardiness Resourcefulness and Emotional Control ($\underline{r}$ = .84), both well above the desired criterion of below .7. Hardiness Resourcefulness, as measured by these items, was not clearly distinguishable from the related areas of Emotional Control, Self Confidence and to some extent Intellectual Flexibility and Productive Teamwork. Perhaps this was reflective of the attempt in this study to combine the notion of 'hardiness' and 'resourcefulness' into a single factor, resulting in a set of items which, although seemingly unidimensional, nevertheless were too closely related to other dimensions of 'life effectiveness' to be considered distinct. Consequently, the Hardiness Resourcefulness items were eliminated and a 36-item, 9-factor model was tested. This produced negligible changes in the fit indices, suggesting that the 'life effectiveness' model fit could still be improved.

One other scale, Productive Teamwork, appeared to be problematic. Both the congeneric and multifactorial CFAs showed two items which loaded below .7 on the latent construct. The MIs were high for Productive Teamwork items on the Social Competence, Task Leadership and Intellectual Flexibility scales. The initial baseline factor correlations of the Productive Teamwork scale with Intellectual Flexibility (r = .73), Task Leadership (r = .74), and Social Competence (r = .69) are moderately high. Substantively it makes sense that being able to function effectively in a team would involve some combination of being socially capable, open to the ideas of others, and having the ability to take on a leadership role when appropriate or needed. This created something of a dilemma -- a meaningful and useful construct, but which was moderately highly correlated with several other scales. Possibly this scale could have been retained for its substantive value, but the desire for developing an instrument that could gather a maximum of information about a person's 'life effectiveness' with minimal items won the day. The Productive Teamwork items were dropped, producing a 32-item, 8-factor model. This improved the fit indices, with the TLI and RNI both over .94.

Although the 32-item, 8 factor model had very good fit indices and was acceptable, it was decided to test whether similar results could be achieved for an even shorter instrument. The weakest item from the eight remaining scales, while checking that the remaining three items in each scale still appeared to represent the original theoretical construct. The resulting 24-item, 8-factor model had a TLI and RNI both over .96.

Insert Table 6 about here

Invariance Testing

An important goal of this research was to develop an instrument that had an equivalent factorial structure for males and females and across a wide range of ages. In order to evaluate success in achieving this goal, tests of the invariance were conducted across
gender and age groups using the 24-item, 8-factor model.

Insert Table 7 about here

Gender

The 24-item 8-factor model separately fitted the male (n = 709) and female (n = 513) groups very well, with TLIs of .960 and .950 respectively (see Table 7). When no invariance constraints were imposed the TLI (.956) was very good, however the imposition of the invariance of factor loading, factor correlations and factor variances actually produced a slightly higher TLI (.957), whereas the added imposition of invariance of uniquenesses led to a small decrement in TLI (.955). Overall, the testing for structural invariance provided very strong support for invariance of the number of factors, factor loadings, factor correlations, and factors variances across males and females, with a TLI of .957. There is also strong support for total invariance with a slightly lower TLI of .955. These results support the conclusion that the LEQ responses measure the same constructs for males and females.

Age

Three age groups were considered: late adolescents (15-18 years, n = 419), young adults (19- 24 years, n = 502) , and adults (25 years and over, n = 302). The 24-item 8-factor model separately fits well for each of these groups, with TLIs of .949, .950, and .937 respectively (see Table 7). When no invariance constraints were imposed the TLI (.946) was very good, however the imposition of the invariance of factor loading, factor correlations and factor variances actually produced a slightly higher TLI (.950), whereas the added imposition of invariance of uniquenesses led to a small decrement in TLI (.945). These results support the conclusion that the LEQ responses measure the same constructs for participants of different ages.

Insert Table 9 about here

Study 2

The 24-item 8-factor model, called the LEQ-H, was administered to a further *960 participants. This allowed for further testing of the LEQ-H model which was desirable, particularly given the substantial reduction in the number of items from initial baseline model.

Congeneric Analyses

Congeneric models with three items are underidentified, that is, they have no degrees of freedom, chi-square equal to zero, and consequently a ‘perfect fit’. One solution to this problem is to reduce the number of parameters estimated by the congeneric model in order to increase the degrees of freedom. This can be done by providing

values for the uniquenesses which, in this case, could be estimated from the Study 1 data. Using this procedure the fit indices reported in Table 10 were produced.

Insert Table 10 about here

The outcome statistics for the one-factor congeneric analyses on the eight LEQ-H scales are presented in Table 3. All scales had co-efficient omegas above .8. Five of the eight scales have TLI/RNIs over the desirable .9, however the Task Leadership TLI/RNI is surprsingly low (.672) and the Time Management TLI/RNI (.870) and the Intellectual Flexibility TLI/RNI (.888).are also below .9. Examination of the factor loadings for these scales nevertheless shows quite high loadings. The cause of these relatively low fit indices appears to have been the estimation of the uniqueness, because for these scales the LISREL output showed high MIs for freeing up these parameters for these three scales. Obviously this technique for dealing with model underidentification is useful when the uniquenesses can be reasonably estimated, but will produce inflated chi-square scores to the extent that the estimated uniquenesses deviate from the uniquenesses of the sample.

Multi-factorial CFA and Invariance Testing

A multifactorial CFA of the 24-item, 8 factor LEQ-H model with 890 cases produced the parameter estimates presented in Table 11 and the factor variances and covariances are in Table 12. This was an excellent model fit ($\chi^2 = 718.94$, $df = 224$) with a TLI of .945 (.020 less than in Study 1) and an RNI of .956 (.016 less than in Study 1). A second-order hierarchical model was also tested, with the 8 scales as first-order factors and a global 'life effectiveness' second-order factor. This model provided a slightly less impressive fit (χ^2=972.55, df =244) with a TLI of .926 and an RNI of .934, but nevertheless can be deemed reasonable support for computation of an additional, global life effectiveness based on a composite of the 8 scale scores. The second-order factor loadings are included in Table 12.

Insert Table 11 about here.

Insert Table 12 about here.

The invariance of the 8-factor LEQ model was tested across males (n = 497) and females (n = 389) and for two age groups, adolescents (11-18 years, n = 677) and adults (19 years and over, n = 210), with the results reported in Table 13.

Insert Table 13 about here.

Gender

The 24-item 8-factor model for the LEQ-H data separately fits the male and female groups well, with TLIs of .939 and .920 respectively (see Table 13). When no invariance constraints were imposed the TLI (.931) was good, however the imposition of the invariance of factor loading, factor correlations and factor variances actually produced a slightly higher TLI (.937), whereas the added imposition of invariance of uniquenesses led to a small decrement in TLI (.935). Overall, this provides very strong support for invariance of the number of factors,

factor loadings, factor correlations, and factors variances across males and females, with a TLI of .937. There is also strong support for total invariance with a slightly lower TLI of .935. These results support the conclusion that responses to the LEQ-H measure the same constructs for males and females.

Age

The 24-item 8-factor LEQ-H model separately fits well for adolescents and adults, with TLIs of .938 and .936 respectively (see Table 13). When no invariance constraints were imposed the TLI (.937) was good. The imposition of the invariance of number of factors and factor loadings produced a slightly higher TLI (.939). The added constraint of invariant factor correlations and factor variances produced a small decrease in the TLI (.936). The final step of modeling invariant uniqueness caused a marked drop in the TLI (.892). Overall these results give very strong support for the invariance of number of factors and factor loadings across the two age groups. There is also strong support for the invariance of factor correlations and factor variances with only a small drop in the TLI. It would appear, however, that the two age groups respond to the items in such a way as to produce different item uniquenesses. Although this differs from the finding in Study 1, it is likely to be caused by the inclusion of a larger proportion of younger participants in the Study 2 sample, particularly school students aged between 14 and 16 years. The important finding, however, is that there is support for factorial invariance and that results from different aged groups' responses to LEQ-H can be meaningfully compared.

Invariance Testing between Study 1 and Study 2

Finally, the data from both studies was pooled to test for invariance of the 24-item 8-factor model across Study 1, which used the 64-item LEQ-G, and Study 2, which used the 24-item LEQ-H (see Table 14). This model fits each study very well, as indicated by the TLIs, for Study 1 (.964) and Study 2 (.945). With no invariance constraints, the fit is also very good (TLI=.957). Imposing of invariance for the factor loadings increases the TLI slightly (.958). Further invariant constraints for factor variances and factor correlations shows a slight drop in the TLI (.956), but this is strong support for invariance across the studies. Total invariance, however, is not supported because when the uniquenesses are constrained to be invariant between the two studies, the TLI drops substantially (.892) -- differences in the age groups involved in each study, as already described, may account for this drop. Overall, these results provide very strong support for the factorial invariance of the 24 common items between the LEQ-G and LEQ-H.

Descriptive Statistics

The strong evidence for factorial invariance justified the pooling of data from Study 1 and Study 2. Descriptives statistics for males, females, and the three age groups based on the pooled data is presented in Table 15, with scale internal consistencies and test-retest correlations in Table 16. Overall, responses to the items are negatively skewed, however the average item mean is low enough (6.13 on an 8 point scale) to allow for more positive self-perceptions to be detected. Internal consistencies are generally satisfactory, particularly given that there are only three items per scale. The test-retest correlations are lower than desirable, however those reported in Table 16 are likely to have been effected, to an unknown degree, by the

intervention period. Ideally further testing of the test-retest stability of the instrument should be conducted.

Summary and Conclusion

The present study had pragmatic, substantive, and methodological aims. The pragmatic purpose was to produce a practical and easy-to-use self-report instrument for evaluating the outcomes of programs which aim to enhance multiple areas of personal effectiveness. As a result, the Life Effectiveness Questionnaires offers a short, reliable, easy-to-use, freely available instrument which has been developed and tested in a wide variety of outdoor education programs and is relevant to the personal development aims which are typical of these programs. Whilst the 24-item, 8-factor can be confidently recommended, it is important to check whether or not the scales are adequate operationalisations of individual program aims and objectives. Conceivably the LEQ could be expanded to include other scales or some scales could be dropped if they are not relevant. It is also possible that alternative factor structures could emerge, depending on the goals of the investigation. Perhaps, for example a hierarchical structure could be further investigated.

Invariance testing indicated that the eight LEQ scales have similar relations with the variables they measure for participants of both genders and three different age groups. This has important implications for researchers and practitioners of personal development programs because the LEQ instrument can be used to make comparative evaluations of program outcomes between gender and age subgroups.

The study's substantive aim was to identify and empirically explore key domains of personal effectiveness which are theoretically amenable to developmental change in experience-based educational and intervention programs. A number of putative scales were tested during pilot studies. Through congeneric and multifactorial CFAs, the 24-item, 8-factor LEQ model was developed. Organisational Self-Discipline had a very high correlation with Time Management and was dropped, while the Hardiness Resourcefulness and Productive Teamwork scales were also dropped because they did not appear to fit single factor congeneric analyses well and cross-loaded onto other factors in multidimensional CFAs. The other eight scales provide unique information about different aspects of life effectiveness and are moderately intercorrelated. A criticism that has been leveled at many CFA studies is that the researcher(s) rarely acknowledge that the fit of favoured models is often identical to other equivalent models (Breckler, 1990). In selecting the 24-item LEQ, it is acknowledged that other statistically defendable models could have been chosen. Researchers who do not share the pragmatic criterion for a quick and simple self-report tool for applied field settings may wish to adopt the 32-item or 40-item LEQ versions, with 4 and 5 items per factor respectively, for which acceptable fit indices over .90 were also obtained. Other researchers may wish to further examine and develop the Organisational Self-Discipline, Productive Teamwork and Hardiness Resourcefulness scales which were dropped from the final model, or may see the need to include additional life effectiveness scales.

The methodological aim was to demonstrate the use of CFA for instrument development. The three step procedure, from congeneric analyses, to multifactorial

analyses, then to invariance testing represents a particularly powerful statistical process. The model developed through this process was strongly supported in the replication study.

Researchers and practitioners aware of the limitations of evaluating the effects of their programs using inappropriate off-the-shelf instrumentation stand to stand to benefit greatly from the development of the LEQ. Further testing of the LEQ in applied settings is recommended. Ultimately, the worth of such an instrument is to examine the effects of intervention programs on participants' 'life effectiveness' and then use these results to proactively improve program design and facilitation.

Table 1. Outline of Life Effectiveness Questionnaire development.

Stage	Version	Year	No. of items / No. of factors	Scale Range	*N*	Description of sample and age range
1	Actuals & Aspirations	1986	38 / 7	1-10	105	OB[a] Standard, Adult, and Management program participants (16 years and over)
2	LEQ-C	1987	72 / 8	1-9	314	OB Standard and Adult program participants (16 years and over)
3	LEQ-D	1987-1988	53 / 11	1-8	335	OB Standard, Adult, and Management program participants (16 years and over)
4	LEQ-E	1989	63 / 7	1-8	114	OB Standard, Adult, and Management program participants (16 years and over)
5	LEQ-F	1990-1993	42 / 8	1-8	406	OB Standard, Adult, Management programs, Army Officer Cadets, Marine Officer Trainees, YE[b] voyages, ANZSES[c] expeditions (16 years and over)
6	LEQ-G	1993-1995	64 / 11	1-8	1164	OB Standard, Adult, and Management programs, YE[b] voyages, ANZSES[c] expeditions, Leeuwin STS[d] , TOE[e] program (15 years and over)
7	LEQ-H	1996-1997	24 / 8	1-8	925	OB Standard, Management, Adventure and School programs, plus OB Staff Applicants and OB Staff, TOE[e] program (11 years and over)

Note: [a]OB = Outward Bound; [b]YE = Young Endeavour Sail Training Scheme; [c]ANZSES = Australian and New Zealand Scientific Exploring Society; [d]Leeuwin Sail Training School; [e]The Outdoor Experience Program

Table 2. Description of the Life Effectiveness Questionnaire dimensions.

LEQ Dimensions	Description
Achievement Motivation[#]	The extent to which the individual is motivated to achieve excellence and put the required effort into action to attain it.
Active Initiative[# *]	The extent to which the individual likes to initiate action in new situations.
Emotional Control[#]	The extent to which the individual perceives he/she maintains emotional control when he/she is faced with potentially stressful situations.
Hardiness Resourcefulness[*]	The extent to which the individual perceives that he/she can perform difficult tasks with limited resources.
Intellectual Flexibility[#]	The extent to which the individual perceives he/she can adapt his/her thinking and accommodate new information from changing conditions and different perspectives.
Organisational Self-Discipline	The extent to which the individual perceives that he/she controls his/her behaviour in a logical, sensible, and organised manner.
Productive Teamwork	The extent to which the individual perceives that he/she has the interpersonal skills to work productively in a team situation.
Self Confidence[# *]	The degree of confidence the individual has in his/her abilities and the success of their actions.
Social Competence[#]	The degree of personal confidence and self-perceived ability in social interactions.
Task Leadership[#]	The extent to which the individual perceives he/she can lead other people effectively when a task needs to be done and productivity is the primary requirement.
Time Management[#]	The extent that an individual perceives that he/she makes optimum use of time.

Note: [*] Dimensions which were introduced for the first time in version G.
[#] Dimensions retained after development of the model in the present study.

Table 3. Factor loadings (FL), uniquenesses, co-efficient (Ω), and goodness-of-fit indices (TLI & RNI) for fitted one-factor congeneric models for the LEQ-G scales.

Scale and Items	FL	Uniqueness	Ω	χ^2	*df*	TLI	RNI
Time Management			.917	76.72	14	.980	.987
TM01 - plan and use my time efficiently	.768	.410					
TM09 - make the most of my time	.750	.438					
TM17 - do not waste time	.812	.340					
TM25 - manage the way I use my time	.886	.216					
TM33 - punctual and do things on time	.588	.655					
TM41 - get information and facts	.550	.697					
TM42 - active person and use time well	.806	.350					
Social Competence			.912	337.78	5	.984	.924
SO02 - successful in social situations	.852	.275					
SO10 - competent in social situations	.857	.265					
SO18 - communicate well with people	.787	.381					
SO26 - confident speaking up	.785	.384					
SO34 - confident speaking to others	.794	.369					
Achievement Motivation			.899	48.83	5	.974	.987
AM03 - get details right	.691	.522					
AM11 - get the best results	.858	.265					
AM19 - do the best I can	.844	.288					
AM27 - do it well	.772	.403					
AM35 - best of every opportunity	.696	.515					
Intellectual Flexibility			.881	66.87	5	.958	.979
IF04 - adapt thinking when plans change	.659	.566					
IF12 - open to ideas and opinions of others	.730	.467					
IF20 - change thinking if a better idea	.692	.522					
IF28 - open to new ideas	.842	.292					
IF36 - adaptable/flexible in thinking/ideas	.828	.315					
Organisational Self-Discipline			.829	243.62	5	.791	.896
OD05 - well organised using reason/facts	.688	.526					
OD13 - personal organisation good	.677	.542					
OD21 - good sense/reason in decisions	.670	.552					
OD29 - plan well & not leave to change	.768	.410					
OD37 - do not put things off	.677	.542					
Productive Teamwork			.881	151.48	5	.895	.948
PT06 - generous and helpful to others	.619	.616					
PT14 - effective team member	.756	.429					
PT22 - co-operate well with people	.830	.311					
PT30 - work well with people	.857	.265					
PT38 - consider the feelings of others	.552	.695					
Task Leadership			.905	33.23	5	.984	.992
TL07 - firm/assertive to get job done	.714	.490					
TL15 - get people to work for me	.809	.345					
TL23 - give clear instructions	.680	.537					
TL31 - good leader	.876	.232					

TL39 - motivate other people	.843	.290					
Emotional Control			.909	44.25	5	.979	.989
EC08 - calm in stressful situations	.841	.292					
EC16 - calm in new situations	.823	.322					
EC24 - calm if people difficult	.632	.600					
EC32 - generally calm & relaxed	.763	.417					
EC40 - calm when things go wrong	.879	.227					
Active Initiative			.928	263.18	14	.936	.957
AI44 - busy & active	.823	.323					
AI46 - getting into things	.809	.346					
AI50 - active & energetic	.786	.382					
AI53 - active in new situations	.776	.398					
AI56 - get action started	.672	.549					
AI59 - active, ‘get into it’ person	.887	.214					
AI62 - make things happen	.742	.449					
Self Confidence			.919	323.69	14	.914	.943
SC45 - ability to do anything I want	.738	.456					
SC48 - if apply self, confident of success	.824	.321					
SC52 - confident of success in action	.850	.278					
SC54 - believe I can do it	.828	.315					
SC57 - confident in decisions I make	.780	.392					
SC60 - trust in myself	.746	.444					
SC64 - in control of my destiny	.559	.688					
Hardiness Resourcefulness			.925	379.00	20	.921	.944
HR43 - positive when things go wrong	.770	.407					
HR47 - cope with difficult situations	.821	.326					
HR49 - perform with limited resources	.757	.426					
HR51 - positive aspects of new situations	.723	.477					
HR55 - resourceful in new situations	.815	.335					
HR58 - efficient in difficult situations	.845	.286					
HR61 - work through long, difficult tasks	.684	.532					
HR63 - persevere when things difficult	.700	.510					

Table 4. Parameter estimates and uniquenesses for the initial baseline 11-factor, 64-item LEQ model.

	Factor Loadings											
Variables[a]	1	2	3	4	5	6	7	8	9	10	11	Uniqueness
AI44	.81	0	0	0	0	0	0	0	0	0	0	.34
AI46	.81	0	0	0	0	0	0	0	0	0	0	.34
AI50	.76	0	0	0	0	0	0	0	0	0	0	.42
AI53	.79	0	0	0	0	0	0	0	0	0	0	.38
AI56	.71	0	0	0	0	0	0	0	0	0	0	.50
AI59	.86	0	0	0	0	0	0	0	0	0	0	.25
AI62	.76	0	0	0	0	0	0	0	0	0	0	.42
AM03	0	.67	0	0	0	0	0	0	0	0	0	.55
AM11	0	.82	0	0	0	0	0	0	0	0	0	.33
AM19	0	.82	0	0	0	0	0	0	0	0	0	.32
AM27	0	.80	0	0	0	0	0	0	0	0	0	.36
AM35	0	.74	0	0	0	0	0	0	0	0	0	.45
EC08	0	0	.83	0	0	0	0	0	0	0	0	.30
EC16	0	0	.84	0	0	0	0	0	0	0	0	.30
EC24	0	0	.65	0	0	0	0	0	0	0	0	.58
EC32	0	0	.78	0	0	0	0	0	0	0	0	.40
EC40	0	0	.88	0	0	0	0	0	0	0	0	.23
HR43	0	0	0	.79	0	0	0	0	0	0	0	.37
HR47	0	0	0	.83	0	0	0	0	0	0	0	.31
HR49	0	0	0	.75	0	0	0	0	0	0	0	.44
HR51	0	0	0	.74	0	0	0	0	0	0	0	.45
HR55	0	0	0	.80	0	0	0	0	0	0	0	.36
HR58	0	0	0	.84	0	0	0	0	0	0	0	.30
HR61	0	0	0	.70	0	0	0	0	0	0	0	.54
HR63	0	0	0	.69	0	0	0	0	0	0	0	.53
IF04	0	0	0	0	.68	0	0	0	0	0	0	.54
IF12	0	0	0	0	.72	0	0	0	0	0	0	.48
IF20	0	0	0	0	.69	0	0	0	0	0	0	.53
IF28	0	0	0	0	.83	0	0	0	0	0	0	.30
IF36	0	0	0	0	.83	0	0	0	0	0	0	.31
OD05	0	0	0	0	0	.65	0	0	0	0	0	.58
OD13	0	0	0	0	0	.71	0	0	0	0	0	.50
OD21	0	0	0	0	0	.65	0	0	0	0	0	.58
OD29	0	0	0	0	0	.72	0	0	0	0	0	.48
OD37	0	0	0	0	0	.74	0	0	0	0	0	.45
PT06	0	0	0	0	0	0	.62	0	0	0	0	.62
PT14	0	0	0	0	0	0	.76	0	0	0	0	.42
PT22	0	0	0	0	0	0	.82	0	0	0	0	.33
PT30	0	0	0	0	0	0	.85	0	0	0	0	.28
PT38	0	0	0	0	0	0	.55	0	0	0	0	.70
SC45	0	0	0	0	0	0	0	.72	0	0	0	.49
SC48	0	0	0	0	0	0	0	.82	0	0	0	.33

SC52	0	0	0	0	0	0	0	.86	0	0	0	.26
SC54	0	0	0	0	0	0	0	.82	0	0	0	.33
SC57	0	0	0	0	0	0	0	.80	0	0	0	.36
SC60	0	0	0	0	0	0	0	.75	0	0	0	.44
SC64	0	0	0	0	0	0	0	.54	0	0	0	.70
SO02	0	0	0	0	0	0	0	0	.84	0	0	.29
SO10	0	0	0	0	0	0	0	0	.85	0	0	.27
SO18	0	0	0	0	0	0	0	0	.80	0	0	.36
SO26	0	0	0	0	0	0	0	0	.81	0	0	.35
SO34	0	0	0	0	0	0	0	0	.81	0	0	.34
TL07	0	0	0	0	0	0	0	0	0	.72	0	.48
TL15	0	0	0	0	0	0	0	0	0	.80	0	.36
TL23	0	0	0	0	0	0	0	0	0	.70	0	.50
TL31	0	0	0	0	0	0	0	0	0	.88	0	.23
TL39	0	0	0	0	0	0	0	0	0	.84	0	.29
TM01	0	0	0	0	0	0	0	0	0	0	.77	.40
TM09	0	0	0	0	0	0	0	0	0	0	.75	.44
TM17	0	0	0	0	0	0	0	0	0	0	.81	.34
TM25	0	0	0	0	0	0	0	0	0	0	.87	.24
TM33	0	0	0	0	0	0	0	0	0	0	.60	.64
TM41	0	0	0	0	0	0	0	0	0	0	.60	.64
TM42	0	0	0	0	0	0	0	0	0	0	.81	.34

Note: [a] AI = Active Initiative; AM = Achievement Motivation; EC = Emotional Control; HR = Hardiness Resourcefulness; IF=Intellectual Flexibility; OD = Organisational Self-Discipline; PT = Productive Teamwork; SC = Self Confidence; SO = Social Competence; TL = Task Leadership; TM = Time Management.

Table 5. Factor variances and covariances for the initial baseline 11-factor, 64-item LEQ model.

	AI	AM	EC	HR	IF	OD	PT	SC	SO	TL	TM
Active Initiative	1										
Achievement Motivation	.70	1									
Emotional Control	.50	.50	1								
Hardiness Resourcefulness	.76	.70	.84	1							
Intellectual Flexibility	.60	.61	.66	.72	1						
Organisational Self-Discipline	.60	.77	.55	.71	.57	1					
Productive Teamwork	.65	.63	.57	.66	.73	.58	1				
Self Confidence	.74	.68	.64	.86	.59	.68	.60	1			
Social Competence	.58	.50	.50	.59	.44	.48	.70	.61	1		
Task Leadership	.68	.60	.60	.74	.52	.64	.69	.74	.73	1	
Time Management	.62	.70	.47	.62	.45	.95	.52	.61	.46	.56	1

Note: [a] AI = Active Initiative; AM = Achievement Motivation; EC = Emotional Control; HR = Hardiness Resourcefulness; IF=Intellectual Flexibility; OD = Organisational Self-Discipline; PT = Productive Teamwork; SC = Self Confidence; SO = Social Competence; TL = Task Leadership; TM = Time Management.

Table 6. Goodness of fit for separate solutions for with no invariance constraints.

Model	χ^2	*df*	TLI	RNI
64-item 11-factor model	11651.95	1897	.8175	.8282
55-item 11-factor model[a]	6540.39	1375	.8826	.8913
50-item 10 factor model[b]	5209.91	1130	.8978	.9058
40-item 10 factor model[c]	3001.80	695	.9244	.9362
36-item 9 factor model[d]	2489.56	558	.9279	.9362
32-item 8 factor model[e]	1927.97	436	.9371	.9447
24-item 8 factor model[f]	771.15	224	.9645	.9712

Note: TLI - Tucker-Lewis index; RNI = relative non-centrality index

[a] 1st item per factor dropped; [b] Organisational Self-Discipline items dropped; [c] 2nd item per factor dropped; [d] Hardiness Resourcefulness items dropped; [e] Productive Teamwork items dropped; [f] 3rd item per factor dropped

Table 7. Goodness of fit for separate solutions with selected invariance contrasts imposed across gender and age groups for the 24-item 8-factor LEQ model.

Model	χ^2	*df*	TLI
No invariance			
Males	595.75	224	.960
Females	528.59	224	.950
15 to 18 years	485.13	224	.949
19 years to 24 years	548.87	224	.950
25 years and over	484.62	224	.937
Gender Invariance			
Null	19626.12	552	.000
No invariance	1124.33	448	.956
NF and FL invariant	1140.86	464	.958
NF, FL, FCr and FV invariant	1243.93	500	.957
NF, FL, FCr, FV and U invariant	1339.15	524	.955
Age Invariance			
Null	20183.78	828	.000
No invariance	1518.63	672	.946
NF and FL invariant	1566.86	704	.948
NF, FL, FCr and FV invariant	1680.20	776	.950
NF, FL, FCr, FV and U invariant	1881.57	824	945

Note: TLI - Tucker-Lewis index; NF = number of factors; FL = factor loadings; FCr = factor correlations; FV = factor variances; U = uniquenesses. The relative noncentrality index (RNI) and the parsimony index (based on the relative noncentrality index) (PRNI) can be calculated from the values presented here. Interpretations based on the RNI are consistent with those based on the TLI. Use of the PRNI would lead to selection of the total invariance models for both gender and age.

Table 9. Internal Consistency and Test-Retest Correlations for the 8-factor, 24-item LEQ-G model.

LEQ-G Scale	Original Model		5 items per scale		4 items per scale		3 items per scale	
	Alpha	Test-Retest *r*	Alpha	Test-Retest *r*	Alpha	Test-Retest *r*	Alpha	Test-Retest *r*
Achievement Motivation	.88	.74	.88	.74	.87	.72	.84	.72
Active Initiative	.92	.74	.91	.74	.90	.73	.88	.73
Emotional Control	.89	.78	.89	.78	.90	.77	.88	.75
Hardiness Resourcefulness	.92	.76	.88	.74	.86	.72	-	-
Intellectual Flexibility	.86	.65	.86	.65	.85	.64	.83	.59
Organisational Self-discipline	.83	.74	.83	.74	-	-	-	-
Productive Teamwork	.85	.73	.85	.73	.82	.70	-	-
Self Confidence	.90	.77	.90	.76	.89	.75	.86	.74
Social Competence	.91	.78	.91	.78	.89	.76	.88	.74
Task Leadership	.89	.81	.89	.81	.88	.81	.88	.81
Time Management	.89	.78	.89	.76	.89	.76	.87	.74
N=1229								

Note: This is a low estimate of the test-retest correlation because the retest was conducted at the completion of Outward Bound or similar programs which is expected to have affected responses.

Table 10. Factor loadings (FL), uniquenesses, co-efficient omega (Ω), and goodness-of-fit indices (TLI & RNI) for fitted one-factor congeneric models for the LEQ-H scales using uniqueness estimates from Study 1 data.

Scale and Items	FL	Uniqueness	Ω	χ^2	*df*	TLI/ RNI	*N*
Time Management			.868	118.20	3	.870	923
TM01 - plan and use my time efficiently	.736	.396					
TM17 - do not waste time	.748	.320					
TM25 - manage the way I use my time	.860	.214					
Social Competence			.896	82.15	3	.929	918
SO02 - successful in social situations	.875	.196					
SO10 - competent in social situations	.865	.202					
SO18 - communicate well with people	.680	.456					
Achievement Motivation			.870	69.12	3	.922	923
AM03 - get details right	.648	.519					
AM11 - get the best results	.887	.185					
AM19 - do the best I can	.755	.350					
Intellectual Flexibility			.825	76.78	3	.888	919
IF20 - change thinking if a better idea	.561	.507					
IF28 - open to new ideas	.813	.299					

IF36 - adaptable/flexible in thinking/ideas	.791	.333					
Task Leadership			.873	268.48	3	.672	914
TL15 - get people to work for me	.630	.347					
TL31 - good leader	.862	.204					
TL39 - motivate other people	.798	.310					
Emotional Control			.883	33.06	3	.977	924
EC08 - calm in stressful situations	.853	.244					
EC16 - calm in new situations	.777	.333					
EC40 - calm when things go wrong	.850	.263					
Active Initiative			.890	40.16	3	.971	922
AI44 - busy & active	.753	.356					
AI50 - active & energetic	.830	.298					
AI59 - active, 'get into it' person	.889	.188					
Self Confidence			.861	38.14	3	.965	925
SC45 - ability to do anything I want	.783	.336					
SC48 - if apply self, confident of success	.752	.371					
SC54 - believe I can do it	.847	.254					

Table 11. Parameter estimates and uniquenesses for the 8-factor, 24-item LEQ-H model.

	Factor Loadings								
Variables[a]	1	2	3	4	5	6	7	8	Uniqueness
AI44	.72	0	0	0	0	0	0	0	.48
AI50	.81	0	0	0	0	0	0	0	.34
AI59	.89	0	0	0	0	0	0	0	.20
AM03	0	.67	0	0	0	0	0	0	.54
AM11	0	.78	0	0	0	0	0	0	.40
AM19	0	.79	0	0	0	0	0	0	.37
EC08	0	0	.81	0	0	0	0	0	.34
EC16	0	0	.80	0	0	0	0	0	.36
EC40	0	0	.85	0	0	0	0	0	.27
IF20	0	0	0	.50	0	0	0	0	.75
IF28	0	0	0	.77	0	0	0	0	.40
IF36	0	0	0	.81	0	0	0	0	.34
SC45	0	0	0	0	.75	0	0	0	.43
SC48	0	0	0	0	.76	0	0	0	.42
SC54	0	0	0	0	.84	0	0	0	.30
SO02	0	0	0	0	0	.85	0	0	.28
SO10	0	0	0	0	0	.85	0	0	.28
SO18	0	0	0	0	0	.71	0	0	.49
TL15	0	0	0	0	0	0	.59	0	.65
TL31	0	0	0	0	0	0	.80	0	.37
TL39	0	0	0	0	0	0	.83	0	.32
TM01	0	0	0	0	0	0	0	.76	.43
TM17	0	0	0	0	0	0	0	.71	.50

TM25	0	0	0	0	0	0	0	.83	.32

Note: [a] AI=Active Initiative; AM=Achievement Motivation; EC=Emotional Control; IF=Intellectual Flexibility; SC=Self Confidence; SO=Social Competence; TL=Task Leadership; TM=Time Management.

Table 12. Factor variances and covariances for the 8-factor, 24-item LEQ-H model and second-order factor loadings.

	AI	AM	EC	IF	SC	SO	TL	TM	2nd-order factor loading
Active Initiative	1								.81
Achievement Motivation	.61	1							.75
Emotional Control	.51	.40	1						.65
Intellectual Flexibility	.48	.62	.54	1					.64
Self Confidence	.77	.65	.59	.50	1				.87
Social Competence	.58	.49	.41	.42	.58	1			.71
Task Leadership	.60	.51	.56	.49	.67	.71	1		.78
Time Management	.58	.71	.48	.43	.64	.50	.58	1	.75

Note: [a] AI=Active Initiative; AM=Achievement Motivation; EC=Emotional Control; IF=Intellectual Flexibility; SC=Self Confidence; SO=Social Competence; TL=Task Leadership; TM=Time Management.

Table 13. Goodness of fit for LEQ-H with selected invariance contrasts imposed across gender and age groups for the 24-item 8-factor LEQ-H model.

Model	χ^2	*df*	TLI
No invariance			
Males	540.41	224	.939
Females	535.10	224	.920
11 to 18 years	614.27	224	.938
19 years and over	423.02	224	.936
Gender Invariance			
Null	11772.51	552	.000
No invariance	1075.51	448	.931
NF and FL invariant	1089.59	464	.934
NF, FL, FCr and FV invariant	1141.93	500	.937
NF, FL, FCr, FV and U invariant	1212.28	524	.935
Age Invariance			
Null	12125.39	552	.000
No invariance	1037.28	448	.937
NF and FL invariant	1054.16	464	.939
NF, FL, FCr and FV invariant	1166.07	500	.936
NF, FL, FCr, FV and U invariant	1706.54	524	892

Note: TLI - Tucker-Lewis index; NF = number of factors; FL = factor loadings; FCr = factor correlations; FV = factor variances; U = uniquenesses. The relative noncentrality index (RNI) and the parsimony index (based on the relative noncentrality index) (PRNI) can be calculated from the values presented here. Interpretations based on the RNI are consistent with those based on the TLI. *check* Use of the PRNI would lead to selection of the total invariance models for both gender and age.

Table 14. Goodness of fit for 24-item 8-factor model with selected invariance contrasts imposed across Study 1 and Study 2.

Model	χ^2	*df*	TLI
No invariance			
Study 1	771.15	224	.964
Study 2	718.94	224	.945
Study Invariance			
Null	30610.25	552	.000
No invariance	1490.09	448	.957
NF and FL invariant	1516.75	464	.958
NF, FL, FCr and FV invariant	1685.89	500	.956
NF, FL, FCr, FV and U invariant	3590.51	524	.892

Note: TLI - Tucker-Lewis index; NF = number of factors; FL = factor loadings; FCr = factor correlations; FV = factor variances; U = uniquenesses. The relative noncentrality index (RNI) and the parsimony index (based on the relative noncentrality index) (PRNI) can be calculated from the values presented here. Interpretations based on the RNI are consistent with those based on the TLI. Use of the PRNI would lead to selection of the NF, FL, FCr, & FV invariance model.

Table 15. Means and standard deviations for the 8-factor, 24-item model for all Study 1 and Study 2 participants, males, females, under 19 years, 19 to 24 years, and 25 years and over.

Item	Total		Males		Females		11 to 18 years		19 to 24 years		25 years and over	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
TM01 - plan and use my time efficiently	5.61	1.44	5.64	1.47	5.58	1.41	5.61	1.47	5.53	1.45	5.73	1.35
TM17 - do not waste time	5.22	1.64	5.25	1.63	5.19	1.65	5.16	1.70	5.16	1.60	5.48	1.51
TM25 - manage the way I use my time	5.54	1.46	5.53	1.48	5.55	1.44	5.51	1.49	5.51	1.44	5.62	1.40
SO02 - successful in social situations	5.70	1.37	5.65	1.39	5.77	1.35	5.73	1.38	5.70	1.35	5.60	1.38
SO10 - competent in social situations	5.85	1.37	5.80	1.38	5.91	1.35	5.83	1.38	5.89	1.35	5.84	1.37
SO18 - communicate well with people	6.10	1.35	5.99	1.38	6.24	1.30	6.12	1.39	6.08	1.35	6.04	1.25
AM03 - get details	6.39	1.30	6.39	1.29	6.39	1.29	6.31	1.39	6.42	1.19	6.56	1.16

right												
AM11 - get the best results	6.66	1.18	6.67	1.17	6.66	1.18	6.60	1.29	6.69	1.09	6.80	0.99
AM19 - do the best I can	6.69	1.20	6.67	1.23	6.73	1.16	6.66	1.27	6.66	1.17	6.81	1.06
IF20 - change thinking if a better idea	6.32	1.30	6.34	1.30	6.29	1.28	6.28	1.42	6.37	1.05	6.35	1.12
IF28 - open to new ideas	6.62	1.15	6.60	1.17	6.64	1.12	6.57	1.25	6.67	1.05	6.65	1.00
IF36 - adaptable/flexible in thinking/ideas	6.31	1.17	6.31	1.20	6.30	1.12	6.23	1.24	6.38	1.08	6.37	1.10
TL15 - get people to work for me	5.55	1.56	5.59	1.60	5.49	1.50	5.29	1.71	5.80	1.34	5.86	1.32
TL31 - good leader	5.96	1.47	5.98	1.49	5.94	1.44	5.91	1.53	6.02	1.42	6.02	1.40
TL39 - motivate other people	5.61	1.44	5.64	1.47	5.77	1.36	5.76	1.48	5.84	1.38	5.81	1.31
EC08 - calm in stressful situations	5.93	1.54	6.10	1.52	5.71	1.55	5.82	1.68	6.11	1.34	5.96	1.42
EC16 - calm in new situations	5.96	1.39	6.08	1.36	5.79	1.40	5.91	1.47	6.07	1.28	5.90	1.30
EC40 - calm when things go wrong	5.96	1.46	6.10	1.44	5.77	1.47	5.90	1.56	6.08	1.30	5.91	1.40
AI44 - busy & actively involved	6.50	1.37	6.44	1.40	6.60	1.29	6.45	1.49	6.54	1.24	6.57	1.20
AI50 - active & energetic	6.56	1.39	6.58	1.41	6.53	1.35	6.58	1.48	6.54	1.29	6.51	1.26
AI59 - active, 'get into it' person	6.41	1.43	6.42	1.47	6.40	1.37	6.43	1.52	6.41	1.36	6.34	1.28
SC45 - ability to do anything I want	6.56	1.44	6.62	1.43	6.50	1.45	6.51	1.52	6.73	1.28	6.45	1.44
SC48 - confident of success	6.29	1.34	6.42	1.31	6.11	1.37	6.21	1.41	6.41	1.27	6.30	1.25
SC54 - believe I can do it	6.62	1.36	6.69	1.34	6.53	1.39	6.65	1.43	6.65	1.26	6.47	1.31
Average item	6.13	0.87	6.15	0.88	6.10	0.83	6.08	0.89	6.18	0.83	6.16	0.85
Listwise *N*	2120		1206		902		1096		616		398	

Note: AI = Active Initiative; AM = Achievement Motivation; EC = Emotional Control; IF = Intellectual Flexibility; SC = Self Confidence; SO = Social Competence; TL = Task Leadership; TM = Time Management.

Table 16. Internal Consistency and Test-Retest Correlations for the 8-factor, 24-item model for all Study 1 and Study 2 participants

LEQ 24-item 8-factor model	3 items per scale	
	Test-Retest *r*	Alpha
Achievement Motivation	.68	.87
Active Initiative	.73	.81
Emotional Control	.75	.87
Intellectual Flexibility	.60	.78
Self Confidence	.73	.84
Social Competence	.75	.86
Task Leadership	.81	.82
Time Management	.75	.84
Average Scale	.72	.84
Total Scale	.67	.93
N	2016	2120

Note: This is a low estimate of the test-retest correlation because the retest was conducted at the completion of Outward Bound or similar programs which is expected to have affected responses.