PROJECT K IN BLACK & WHITE

A theory-driven & randomised controlled trial evaluation of a youth development programme

Kelsey Lauren Deane

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psychology

The University of Auckland, 2012
Abstract

Project K is a New Zealand-based positive youth development programme that combines a three-week wilderness adventure, a five week community exploration and service component, and a year-long mentoring partnership to promote positive growth in 14-15 year olds with low self-efficacy. The primary intention for this research was to produce credible knowledge about Project K in a manner that empowered programme stakeholders to address their information needs. Another intention was to produce knowledge that contributes to academic understanding of youth development programming and programme evaluation.

A comprehensive mixed-method evaluation of Project K was guided by a novel theoretical framework, the Thoughtful Evaluation model. The model stresses an approach that balances aspects of scientific credibility and stakeholder empowerment. Findings from eight programme staff focus groups, 361 open-ended comments about the programme from youth participants, four key programme documents, results from six previous Project K research projects and thorough literature reviews of youth development, adventure, service-learning, and mentoring programming were triangulated and integrated into a programme theory of change. A logic model depicting the theory suggests that Project K reinforces participant growth with experiential learning and scaffolded support using an iterative and cyclical process. Characteristics of the individuals involved and the programme environment reportedly influence the success of this process.

Hierarchical linear modelling of data from 1092 Project K and Control participants (at baseline) collected as part of a randomised controlled trial revealed that Project K was effective in improving academic and social self-efficacy from pre to post-programme. Effects were sustained one year post-programme. Project K participants reported higher career decision self-efficacy one year post-programme but no difference existed between the groups for achievement on the major national qualification offered in New Zealand secondary schools (NCEA, Level 1).
The interaction effects obtained between programme condition and gender, ethnicity, or socioeconomic school status indicated that Project K appears to reduce disparities between different participant subgroups over the follow up period.

While limitations were identified with Project K, this research demonstrated that it is a “model” programme in terms of best practice principles and programme effects. The rigorous, comprehensive and process-focused nature of the evaluation offers a valuable contribution to youth development programme designers, practitioners, policymakers, and researchers interested in enhancing the function of similar programmes. In addition, the Thoughtful Evaluation model is offered as a useful vision for single-evaluator situations.
Dedication

For Luis Alberto Vallejas Ramos
Acknowledgements

I feel incredibly fortunate to have been surrounding by many amazing and supportive individuals over the course of this project. The most instrumental to my progress as a budding academic was my phenomenal supervisor, Niki Harré. Her guidance has been exceptional at every step of the journey, and her patience and empathic understanding never-ending. All in all, Niki exhibits every quality you could hope to have in a mentor. Moreover, she is an inspirational role model for living life with integrity, generosity, passion, and commitment. Niki, I cannot thank you enough and look forward to what future collaborations with you will bring.

I want to thank the individuals at the Foundation for Youth Development (Jo-anne Wilkinson and Graeme Dingle especially) for inviting me into their whānau, sharing their experiences, involving me in the development of the organisation, and for continuing to inspire me. To Julie Moore, thank you for your tireless commitment to this project, your enthusiastic involvement in our research collaborations and your faith in my abilities. To Victoria Hood, thank you for involving me in programme development opportunities and for being so encouraging. Thanks also to Priscilla Lai, Alison Mendelew and Stas Popov for accommodating all my requests in such an efficient manner and to everyone else for listening, giving input, and always welcoming me through the door.

I would like to thank the Project K Programme Directors, Mentor Coordinators, programme providers, and other staff for taking the time to share their expertise and provide valuable feedback on the findings. I have so much admiration for all the energy you put into supporting young people in this country.

I would also like to acknowledge Chris Sibley, Heidi Leeson, Margaret Francis, Larissa Isted and Peter Muris for help with either statistical, technical or data entry aspects. Thank you to Shannon Johnston, Kane Meissel, and Nickola Overall for proofreading sections of this thesis and Jade LeGrice for consult on the Māori aspects of the research and writing. I appreciate all of the valuable feedback. To Andrew Lavery at Academic Consulting Ltd, I am so grateful for your exceptional formatting skills. This also would not have been possible without the support of many others in the Department of Psychology nor without financial support from the New Zealand International Doctoral
Research Scholarship and University of Auckland International Doctoral Research Scholarship. In addition, I would like to thank my Honours supervisor at the University of New Brunswick, Daniel Voyer. He helped sow the seeds of my postgraduate career by providing early opportunities to cultivate my research skills and by expressing strong belief in my capabilities.

A special thanks to a few of my wonderful friends/role models/mentors within the Department of Psychology at Auckland University. Nickola Overall, you have had a substantial influence on my development as an academic, my health, and my happiness through your friendship and extensive mentorship. Annette Henderson, Pat Bullen, Jan Trayes, and Shannon Johnston you have all also added so much sunshine to my life since moving to New Zealand.

A number of other people have contributed indirectly to this colossal achievement either by providing necessary support in times of need and/or positivity to my life over the course of the endeavour: Mary-Jo, Amanda, Vicki, Jane, Katie, Meghan, Annie, Ali, Iannick, Des, Julie, Judy, Barney, Rob, Kathy, Sharon, and Michael. To Jarrod, thank you for accompanying me on the majority of this journey. You have shown me how to make extraordinary goals happen and you helped me to realise my passion for youth work and research. Vitali, thank you for being such a supportive partner; not only do you tolerate my emotionality during stressful times but you always find ways to brighten my day.

I give huge thanks for my incredible family (extended and immediate). They have always supported my dreams, despite these continually taking me to far-away lands. The distance has been difficult, especially for my Nana and Pampy, but I appreciate your patience and you making the most of the short periods we do spend together. To Mom, Dad, Kendra, Wes, and Taryn, thank you for consoling, counselling, and supporting me in so many ways throughout the process, especially during the hard bits. Your constant friendship and ability to make me laugh have been so important to my success.

Finally, I would like to give additional thanks to my mother, Ann Deane. She committed so much time and energy to motivating and encouraging me during the most difficult parts of the journey. I know I would not have made it this far without her unyielding support.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>III</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>V</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>VII</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>XIII</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>XV</td>
</tr>
<tr>
<td><strong>CHAPTER ONE</strong>  THE PROJECT K EVALUATION SCENE PROGRAMME &amp; RESEARCH OVERVIEW</td>
<td>1</td>
</tr>
<tr>
<td>THE SOCIO-POLITICAL CONTEXT OF PROJECT K</td>
<td>1</td>
</tr>
<tr>
<td>THE EVALUAND</td>
<td>1</td>
</tr>
<tr>
<td>THE EVALUATOR’S POSITION</td>
<td>10</td>
</tr>
<tr>
<td>THESIS AIMS &amp; CONTRIBUTIONS</td>
<td>13</td>
</tr>
<tr>
<td><strong>CHAPTER TWO</strong>  THOUGHTFUL EVALUATION THE OVERARCHING FRAMEWORK</td>
<td>15</td>
</tr>
<tr>
<td>EVALUATION THEORY &amp; PRACTICE</td>
<td>15</td>
</tr>
<tr>
<td>THE HISTORY OF EVALUATION AS A DISTINCT DISCIPLINE</td>
<td>16</td>
</tr>
<tr>
<td>THE ESSENTIAL FEATURES</td>
<td>21</td>
</tr>
<tr>
<td>A VISION FOR MY PRACTICE</td>
<td>26</td>
</tr>
<tr>
<td>THE THOUGHTFUL EVALUATION MODEL</td>
<td>27</td>
</tr>
<tr>
<td>CHRONOLOGY OF THE EVALUATION PROCESS</td>
<td>32</td>
</tr>
<tr>
<td><strong>CHAPTER THREE</strong>  POSITIVE YOUTH DEVELOPMENT THROUGH ADVENTURE, SERVICE &amp; MENTORING</td>
<td>39</td>
</tr>
<tr>
<td>ADOLESCENT DEVELOPMENT &amp; YOUTH PROGRAMMING</td>
<td>39</td>
</tr>
<tr>
<td>A POSITIVE AND ECOLOGICAL APPROACH TO DEVELOPMENT</td>
<td>40</td>
</tr>
<tr>
<td>YOUTH DEVELOPMENT IN AOTEAROA NEW ZEALAND</td>
<td>43</td>
</tr>
<tr>
<td>THE VALUE OF YOUTH PROGRAMMES</td>
<td>49</td>
</tr>
<tr>
<td>THE FEATURES OF YOUTH DEVELOPMENT PROGRAMMES</td>
<td>50</td>
</tr>
<tr>
<td>THE STATE OF YOUTH PROGRAMME EVALUATION STUDIES</td>
<td>53</td>
</tr>
<tr>
<td>OUTDOOR ADVENTURE PROGRAMMES</td>
<td>56</td>
</tr>
<tr>
<td>THE ORIGIN AND GROWTH OF OUTDOOR ADVENTURE PROGRAMMING</td>
<td>56</td>
</tr>
<tr>
<td>OUTDOOR EDUCATION, ADVENTURE PROGRAMMES, WILDERNESS THERAPY - ONE AND THE SAME?</td>
<td>57</td>
</tr>
<tr>
<td>RATIONALE FOR ADVENTURE PROGRAMMING</td>
<td>58</td>
</tr>
<tr>
<td>PROCESSES OF CHANGE</td>
<td>59</td>
</tr>
<tr>
<td>THE SUCCESS OF ADVENTURE PROGRAMMES</td>
<td>66</td>
</tr>
</tbody>
</table>
ADVENTURE PROGRAMMES, NATURAL LANDSCAPES, AND THE NEW
ZEALANDER MENTALITY ................................................... 70
DEVELOPMENT THROUGH SERVICE ..................................... 76
THE ALIENATION OF YOUNG PEOPLE FROM COMMUNITY LIFE .... 77
SERVICE-LEARNING AS A SOLUTION .................................... 78
THE CENTRAL ELEMENTS & OPTIMAL CONDITIONS ................. 80
THE THEORETICAL PROCESSES OF SERVICE-LEARNING ........... 84
EVIDENCE OF EFFECTIVENESS IN VIEW OF POPULAR SUPPORT .... 85
PARTICIPANT-ENVIRONMENT INTERACTIONS .......................... 89
MENTORING PROGRAMMES ............................................. 91
THE RATIONALE FOR MENTORING PROGRAMMES .................... 91
MENTORING RELATIONSHIPS ........................................... 93
THE EFFECTIVENESS OF MENTORING PROGRAMMES .............. 94
MENTORING PROCESSES AND CONDITIONS PROMOTING GROWTH 97
MENTORING IN THE NEW ZEALAND CONTEXT ......................... 107
COMMON YOUTH PROGRAMMING THEMES .............................. 111
DISTINCTIVE FEATURES ................................................. 112
THE STATE OF AFFAIRS IN NEW ZEALAND ............................... 113

CHAPTER FOUR  CONSTRUCTION OF THE WHITE BOX PROJECT K’S

THEORY OF CHANGE .......................................................... 115

THEORY-DRIVEN EVALUATION ............................................. 115
CHARACTERISTICS OF A THEORY-DRIVEN APPROACH ............ 116
BENEFITS OF A THEORY-DRIVEN APPROACH ......................... 117
A WHITE BOX EVALUATION OF PROJECT K ........................... 118
THEORETICAL INFLUENCES OF THE CURRENT APPROACH .......... 119
THE CURRENT APPROACH ................................................. 120
METHODOLOGY ............................................................. 120
THE THEORY- BUILDING PROCESS .................................... 120
EXPLORATORY THEORY DEVELOPMENT: GENERATING THE THEORY 124
PRELIMINARY RESULTS ...................................................... 127
THE ANTECEDENT CONDITION ........................................... 128
THE STUDENT PROFILE .................................................. 129
THE ESSENTIAL PROGRAMME STRATEGIES .......................... 130
FACTORS INFLUENCING PROGRAMME SUCCESS ..................... 133
THE KEY PROJECT K OUTCOMES ........................................ 138
SUMMARY OF THE STAKEHOLDER-DRIVEN LOGIC MODEL ........ 143
STEP 2 OF EXPLORATORY THEORY DEVELOPMENT: ASSESSING CONSENSUS ... 144

The Analytic Strategy ....................................................... 144
Comparative Data Source 1: Youth Participant Responses ............ 145
Comparative Data Source 2: Document Review ........................ 150
Comparative Data Source 3: Previous PROJECT K Research ........ 157
DISCUSSION ................................................................. 159
MODEL FIT ACROSS THE PROJECT K DATA SOURCES ........................................... 159
ASSESSING THE SURFACE LEGITIMACY: A DIRECT LOGIC ANALYSIS ............. 165
STUDY LIMITATIONS, CONTRIBUTIONS AND CONCLUSIONS ....................... 175

CHAPTER FIVE EVALUATING THE BLACK BOX PROJECT K’S IMPACT ON
EFFICACY & ACADEMIC ACHIEVEMENT OUTCOMES .................................. 183

THE CENTRALITY OF COMPETENCY-BELIEFS ................................................. 183
EXPECTANCIES, VALUES AND HUMAN POTENTIAL ........................................... 186
SELF-EFFICACY WITHIN ACADEMIC, SOCIAL AND CAREER DEVELOPMENT
SPHERES ............................................................................................................. 191
SELF-EFFICACY ENHANCING INTERVENTIONS ............................................... 195
GROUP DIFFERENCES IN EFFICACY BELIEFS .................................................. 197
THE CURRENT STUDY .......................................................................................... 199

METHODS........................................................................................................... 201

RESULTS.............................................................................................................. 236

DESCRIPTIVE STATISTICS AND BIVARIATE CORRELATIONS ......................... 236
UNCONDITIONAL MODELS AND INTRA-CLASS CORRELATIONS ....................... 243
MAIN EFFECT OF PROJECT K ON DESIRED OUTCOMES ................................ 254
MAIN EFFECTS OF DEMOGRAPHIC VARIABLES ON OUTCOMES .................... 258
MODERATING INFLUENCES OF DEMOGRAPHIC VARIABLES ON PROJECT K
EFFECTS .............................................................................................................. 260
OVERVIEW OF KEY FINDINGS IN PLAIN TERMS ............................................. 269

DISCUSSION ......................................................................................................... 273

OVERVIEW OF STUDY LIMITATIONS AND CAVEATS ..................................... 287

CONCLUSION ....................................................................................................... 296

CHAPTER SIX SELF-REFLECTIVE META-EVALUATION THOUGHTFUL
EVALUATION IN ACTION .................................................................................... 297

THE META-EVALUATIVE MIND-SET ................................................................... 297

CREDIBLE FEATURES OF THE CURRENT RESEARCH ..................................... 298

EMPOWERING FEATURES OF THE CURRENT RESEARCH ................................. 299

SUCCESSES OF THE APPROACH ....................................................................... 301

INTERESTING TENSIONS ................................................................................... 303

CONCLUSION ....................................................................................................... 308

CHAPTER SEVEN FINAL THOUGHTS A RESEARCH REPRISE ......................... 309

ADVANCING KNOWLEDGE BEYOND THE LOCAL SCENE ................................. 311

LIMITATIONS AND LESSONS FOR PROGRAMME DEVELOPMENT ................. 316

CLOSING REMARKS ........................................................................................... 321

APPENDICES ....................................................................................................... 323

Appendix A: Ethical Documentation for Thesis ..................................................... 325

Appendix B: Logic Model Materials .................................................................... 347

Appendix C: Foundation for Youth Development Ethical Documentation for Project
K Randomised Controlled Trial Evaluation .......................................................... 353
Appendix D: Foundation for Youth Development Self-Efficacy Questionnaires.............. 373
Appendix E: Supplementary Analyses for Chapter Five Randomised Controlled Trial Study .............................................................................................................................................. 390
Appendix F: Example Outputs from Stakeholder Empowering Processes..................... 401
REFERENCES ........................................................................................................................................................................ 411
Table 1. The essential features associated with high quality evaluation practice......................... 24
Table 2. Types of New and Challenging Experiences Provided by Project K................................ 131
Table 3. Immediate and Short-Term Outcomes with Associated Quotes ................................ 140
Table 4. Outcomes identified by Project K participants ............................................................... 148
Table 5. Project K Document Review Findings........................................................................ 152
Table 6. Summary of Results of the Direct Logic Analysis ..................................................... 166
Table 7. Comparisons of actual and parallel analysis eigenvalues for self-efficacy items at Time 1 (specifying a two factor extraction)........................................................................ 211
Table 8. Comparisons of the actual and parallel analysis eigenvalues for self-efficacy Items at Time 3 (specifying a 3 factor extraction)........................................................................ 212
Table 9. Comparisons of baseline frequencies and percentages for breakdown of ethnic categories across Project K and Control conditions using both the Single/Combination and Total Response output strategies......................................................................................... 217
Table 10. Frequency and percentages for participant non-inclusion......................................... 220
Table 11. Descriptive statistics for Level-1 continuous variables ............................................. 239
Table 12. Standardised mean difference scores between Project K and Control accounting for baseline scores ......................................................................................................................... 239
Table 13. Bivariate Correlations between Level-1 Variables..................................................... 242
Table 14. Multilevel modelling results for academic self-efficacy linear and curvilinear growth trajectories......................................................................................................................... 246
Table 15. Multilevel modelling results for social self-efficacy linear and curvilinear growth trajectories............................................................................................................................. 247
Table 16. Variance components and intra-class correlation coefficients for three-level unconditional models ........................................................................................................................... 248
Table 17. Variance components and intra-class correlation coefficients for two-level models ..... 248
Table 18. Multilevel model results for post-programme academic and social self-efficacy........... 249
Table 19. Multilevel model results for one year post-programme academic self-efficacy............. 250
Table 20. Multilevel model results for one year post-programme social self-efficacy................ 251
Table 21. Multilevel model results for one year post-programme career-decision self-efficacy and NCEA Level 1 Credits ........................................................................................................ 252
Table 22. Multilevel model results for one year post-programme NCEA Level 1 Credits ........... 253
Table 23. Empowering Evaluation Processes............................................................................. 300
**List of Figures**

*Figure 1.* The Thoughtful Evaluation Model illustrating the role of the evaluator in balancing the credible and empowering aspects of evaluations to meet the needs of external and internal programme stakeholders. ................................................................. 28

*Figure 2.* The chronology of the research process illustrating the concurrence of the two core thesis studies, the order of the various literature reviews, and the on-going nature of the stakeholder consultation and evaluator self-reflection processes. ......................... 34

*Figure 3.* The Programme Theory-Building Process ordered by stages and steps according to Project K data source, method type, and analytic strategy. ......................... 123

*Figure 4.* Preliminary Project K logic model detailing the antecedent condition (purple), characteristics of the target group (blue), key programme processes (yellow), the influencing factors (light green and orange), short-term outcomes (pink) and intermediate outcomes (green).................................................................................. 128

*Figure 5.* The Finalised Project K Logic Model depicting the antecedent condition (purple), characteristics of the target group (blue), key programme processes (yellow), the influencing factors (light green and orange), short-term and intermediate outcomes (pink) and desired long-term outcomes (green). ................................................................. 160

*Figure 6.* Frequency distribution for school decile ratings across programme sites. .......... 216

*Figure 7.* Example of a three-level hierarchical data structure for a multisite evaluation where the full programme may fall under a national governance (like Project K) but programmes at each regional site are allocated to Level-3, individuals within each site are at Level-2, and repeated observations of/from these individuals are at Level-1........ 224

*Figure 8.* Linear growth trajectory of academic self-efficacy for Project K and Control from baseline to one year post-programme. ................................................................. 254

*Figure 9.* Linear growth trajectory of social self-efficacy for Project K and Control from baseline to one year post-programme. ................................................................. 255

*Figure 10.* Curvilinear growth trajectory of academic self-efficacy for Project K and Control from baseline to one year post-programme .................................................. 256

*Figure 11.* Curvilinear growth trajectory of social self-efficacy for Project K and Control from baseline to one year post-programme. ................................................................. 257

*Figure 12.* Interaction between programme condition and gender for academic self-efficacy one year post-programme (sustained programme effect). ........................................ 261

*Figure 13.* Interaction between programme condition and gender for the additional academic self-efficacy follow up effect.................................................................................. 262

*Figure 14.* Interaction between programme condition and Māori ethnicity for social self-efficacy one year post-programme (sustained effect). .................................................. 263

*Figure 15.* Interaction between programme condition and Māori ethnicity for the additional social self-efficacy follow up effect. ................................................................. 264

*Figure 16.* Interaction between programme condition and Asian ethnicity for social self-efficacy one year post-programme. ................................................................. 265

*Figure 17.* Interaction between programme condition and school decile rating for the additional social self-efficacy follow up effect. ................................................................. 266

*Figure 18.* Interaction between programme condition and Asian ethnicity for career decision self-efficacy one year post-programme. ................................................................. 267
Figure 19. Interaction between programme condition and Asian ethnicity for NCEA Level 1 credits one year post-programme................................................................. 268

Figure 20. Interaction between programme condition and school decile rating for NCEA level 1 credits one year post-programme................................................................. 269
Chapter One

THE PROJECT K EVALUATION SCENE
PROGRAMME & RESEARCH OVERVIEW

The practice of evaluation is not separable from the socio-political practices and institutions to which it is designed to contribute or in which it is embedded ~ Jennifer Greene (2011, p.7)

THE SOCIO-POLITICAL CONTEXT OF PROJECT K

Programme evaluation is a politically-charged process. First, every social programme is underpinned by the values of its stakeholders and the socio-political context in which it arose. Second, the evaluator brings another layer of expectations and attempts to judge the worthiness of the endeavour. This thesis is structured around a programme evaluation. Accordingly, it begins by describing the conditions under which the idea for Project K (the programme under investigation or evaluand) arose. I also describe its current operations and the situation surrounding the evaluation of Project K when I came on board as an external evaluation researcher in mid-2007. Then I outline the general perspective that underlies my approach to this piece of research and, consequentially, my interpretations of the findings. To conclude the chapter, I present the overarching thesis aims and contributions.

THE EVALUAND

Program Inception

Project K was launched in 1996 by prominent New Zealand mountaineer Graeme Dingle and his partner, Jo-anne Wilkinson. It was the early 1990s when the pair returned to New Zealand from an expedition in the Arctic feeling perturbed by the dysfunction they had observed in some of the communities they had visited (these were described as having high rates of substance abuse and violence with few opportunities for young people). Dingle and Wilkinson were alarmed to learn,
from frequent media reports, that many young people in New Zealand faced similar obstacles. It was this realisation that prompted action towards finding a solution that would generate better outcomes for young New Zealanders. That solution was Project K\(^1\) (Dingle, 2005; Dingle & Wilkinson, 2011).

The idea for a programme that combined adventure programming with a community-based component and a year of mentoring emerged following extensive research and consultation with educational experts in New Zealand. Initially, the programme was conceptualised as an intervention for high risk youth or what Dingle (2005) described as those “who had already gone off the rails” (p. 237) but Dingle and Wilkinson’s research indicated that many programmes already existed for this demographic whereas there appeared to be a gap in services for those that could easily end up heading in that direction (Dingle, 2005; Dingle & Wilkinson, 2011).

**Programme Description\(^2\)**

In its current state, Project K can be described as a New Zealand-based positive youth development programme that aims to improve the psychological, social and physical well-being of 13 to 15 year old youth identified as having low self-efficacy, but who do not exhibit very high risk behaviour (extensive details of the participant selection process are provided in Chapter Five). To facilitate positive change, Project K utilises a multi-component strategy that incorporates a Wilderness Adventure, a Community Challenge, and one-to-one mentoring. The full programme occurs over the course of 14 months and each programme serves 12 students (ideally with an equal number of males and females).

**The Wilderness Adventure**

Following a two to three day induction period during which participants are oriented to the nature of the programme, they embark on the two (at minimum) but preferably three-week Wilderness Adventure. Participants are

---

\(^1\) The “K” stands for **k**oru, the Māori symbol for new life and sustainability but also for other “K” words Dingle felt were powerful such as **k**iwi, **k**ia ora, and **k**ia kaha (Dingle, 2005).

\(^2\) The programme description is based on my extensive communication with key programme stakeholders and review of relevant programme documents.
transported to an outdoor camp where they prepare for a seven to ten day expedition by developing a number of requisite skills (e.g. goal-setting, problem-solving, communication, and leadership) through participation in teambuilding and challenge-based activities. The expedition brings participants further into the wilderness and may involve a number of water-based and/or land-based activities such as kayaking, sailing, cycling, hiking, abseiling, and short solos (1-2 hours). Each participant takes a turn being a leader for a day and at the end participants must identify goals for the year ahead. At least four goals must be selected and one of these must be of an academic nature whilst another must be related to physical health.

*The Community Challenge*

In the next stage, participants are reintegrated into their day to day environment and encouraged to transfer the lessons learned during the Wilderness Adventure to new community-based challenges. The Project K group reunites either one day a week over 10 weeks, two days a week over 5 weeks or every day for two consecutive weeks immediately upon completing the Wilderness Adventure. In teams, participants organise activities which require the exploration of community facilities and resources and networking with key community figures. Community Challenge facilitators support each individual in identifying his or her future aspirations and taking the initial steps to achieve these. Workshop sessions addressing issues pertinent to adolescent health and well-being (e.g. relationships, sex, drugs and alcohol, nutrition) are provided and participants develop a one to two-day collaborative project designed to address a community need.

*The Adult-Youth Mentoring Partnership*

Adult mentors are acquainted with the Project K group at the beginning of the Community Challenge and soon after each mentor is matched with one youth participant. All mentors undergo a screening and interview process and 20 hours of training before the match takes place. Youth are paired with an adult of the same gender if at all possible. Participants and mentors preferences are taken into account but the young person’s preference is the priority when coordinating the
pairs. Over the next 12 months, the pair must communicate via phone or email each week and meet in person on a bimonthly basis. In addition to taking part in informal fun activities and developing a non-judgmental friendship with the young person, a large focus of the mentoring component is to provide support with the goal-setting process. A mentor coordinator monitors the progress of the relationship and offers support when needed. Group activities for all mentors and mentees are also organised throughout the year. The end of the relationship (and Project K) is marked by a graduation celebration.

Programme Expansion, Centralised Governance & Localised Implementation

Project K was endorsed (and continues to be) by many prestigious New Zealanders from the outset. The first Project K trustees included revered mountaineer and philanthropist Sir Edmund Hillary; former Governor-General Sir Paul Reeves; District court judge Dame Augusta Wallace; Chief Executive Officer of the Australia New Zealand banking group, Steve Jones; high profile politician, the Honourable Fran Wilde; prominent business leader Wayne Walden; and the programme co-founder and well-known mountaineer, Graeme Dingle (Dingle, 2005). The programme’s credibility stems from its well-researched and evidence-based design yet the long-lasting support from esteemed New Zealanders has surely helped with securing continued funding and influenced the programme’s rapid growth, as described below.

Today Project K is a national multi-site programme that operates several times a year across eleven different regions of the country. In 2011, programmes were delivered through 18 participating schools. The programme is now owned by the Foundation for Youth Development (FYD), a youth development organisation co-founded by Dingle and Wilkinson that provides programme governance, materials, and evaluation and research for four different youth programmes through its National Support Office: Stars (a peer mentoring programme for all Year 9 students in participating schools); Kiwi Can (a school-wide values-based programme for students in primary and intermediate schools); Male Youth New Directions or MYND (a mentoring programme for high-risk male young offenders); and Project K (the flagship programme; www.fyd.org.nz).
FYD licenses Community Partners (local trusts) to deliver Project K (and the three other programmes) within a specific region of the country. The programmes are community-driven in the sense that each region has their own board of trustees and staff and volunteers are sourced from local communities. The activities within each component may also be adapted to suit the local culture and environment; however the core structure of the programme is standardised across the regions. FYD ensures quality control by monitoring each programme site’s operations and by providing standardised policies and procedures (i.e. selection criteria, training, research and evaluation and general programme materials; www.fyd.org.nz).

FYD is known as an organisational leader in New Zealand youth development. The programme founders have remained active in the day to day operations of the National Support Office and they have cultivated an organisational climate where the goals of continual improvement and accountability are promoted. As a result, both internal and external research and evaluation of its programmes is welcomed and encouraged.

**Previous Project K Research**

The largest evaluation initiative undertaken by FYD to date is a randomised controlled trial (RCT) outcome evaluation of Project K. It was implemented under direction from the Ministry of Social Development (MSD) to ensure the programme objectives were being met. The evaluation measures and data collection protocols were designed by FYD in collaboration with researchers from the University of Auckland (see Moore, 2005). Individuals from MSD’s Centre for Social Research and Evaluation analysed the data from a subset of eight Ministry-funded programme sites³ (Qiao & McNaught, 2007).

Potential Project K participants in each region were identified through Moore’s (2005) academic, social and help-seeking self-efficacy survey which was delivered to all students in the 2nd year of participating secondary schools. In addition, two teachers each completed a three-item self-efficacy measure for every

³ Throughout this thesis I use the term “programme site” to refer to the group encompassing the 12 Project K and the 12 Control group students from the same school.
student at this grade level. Individuals who consented to participate were then randomly allocated to the programme, a control group, or a reserve group. A health and lifestyle behaviours questionnaire (Moore, 2005) was administered to the programme and control groups prior to programme start. This measure and the self-efficacy survey were re-administered at the end of the programme and one year after programme completion. A career decision self-efficacy subscale was also included in the survey materials at this point (the full details of the randomised controlled trial process are provided in Chapter Five).

Project K staff members were responsible for enacting the data collection procedures and/or training volunteers to assist with the process. This resulted in an extensive database of evaluation information (and it continues to grow as three-year post-programme measures are collected). Several research projects have centred on analyses of some of this data, one of which is a core component of this thesis and detailed in Chapter Five. Project K has also been the focus of several other non-RCT related studies.

To further set the scene prior to moving into the specifics of the current research, I provide a brief overview of the findings from Qiao and McNaught’s (2007) analysis of the MSD-funded subset of RCT data and five other research projects that have particular relevance to the current research. Two of these projects were conducted prior to the Qiao and McNaught study; the three other projects were conducted concurrently with my thesis research but have now been completed.

Qiao & McNaught (2007) found that the academic and social self-efficacy levels of Project K participants improved more so than the control group’s between the pre and post-programme measurements and the differences were maintained one year post-programme. Help-seeking self-efficacy patterns were similar although there was no difference between the groups at one year post-programme. No differences were obtained between the groups on measures of health and lifestyle behaviours (e.g. physical activity, eating and risk behaviours, prosocial activities, etc.). They also found that some programme effects differed according to
participant ethnicity. For example, career-decision self-efficacy was higher for New Zealand European and Māori Project K students one year after completion compared to these control subgroups and Māori Project K students obtained a significantly higher number of NCEA credits than Māori control students one year post-programme.

*Warren (2005)* evaluated changes in the happiness, family cohesion, and nature of goal-directed activities taken on by Project K participants and a non-randomised control group from three New Zealand schools. She evaluated these outcomes prior to programme commencement, after the Community Challenge, and one year after the pre-programme measurement. She also interviewed Project K participants from one school at each of these time points to assess their perceptions of the programme.

She discovered that, on average, the happiness and family cohesion of Project K students increased over the first two stages of the programme and the increase in happiness was maintained when assessed at the end of the programme. In contrast the control group experienced no changes in these variables over time. Project K participants reported an increase in goal-directed activities associated with independence and agency. Their goals also became more closely aligned with their personal values whereas the control group did not experience any goal-related changes. The interviews revealed that participants really enjoyed the Wilderness Adventure but the Community Challenge was boring for most in comparison; the mentoring component was rated very positively overall and participants indicated they gained the most from this component. They reported that mentors helped by motivating them to achieve their goals, by introducing them to new experiences, and by providing companionship. However, when interviewed after the Community Challenge, most participants expressed anxiety that the time away from school during the first two programme stages had compromised their academic progress and explained that teachers did little to help them reintegrate into the classroom environment.
O’Neill (2005) evaluated the 20-hour mentor training programme FYD requires potential mentors to undertake prior to being matched with a Project K participant. The training programme was delivered in 6 modules focusing on: 1) the programme’s history, creating a safe environment, mentoring best practices; 2) contemporary youth issues, building resiliency and efficacy; 3) the characteristics of a positive relationship, exploring personal and Project K values, cultural differences and worldviews; 4) goal-setting (including different cultural and gendered understandings of it), the boundaries of the relationship; 5) issues of confidentiality, working with the mentee’s family; 6) celebrating achievements, self-evaluation and reflection, and the appropriate way to dissolve relationships. O’Neill discovered that this training programme met all of the recommended guidelines for mentor training, as outlined in the international literature. Furthermore, the mentor-trainees scored significantly higher and rated themselves significantly better on relevant skills after the training programme.

Somervell (2011) explored the process implemented by the Community Partners to identify the participants to be excluded from the Project K selection pool due to their high risk status (i.e. participants with a history of violence, substance abuse, suicide attempts, regular counselling or serious problems or severe cognitive/learning difficulties). She discovered that the eligible Project K participants and the excluded students did not differ significantly in their self-efficacy levels. She also noted that the way in which the exclusion criteria was employed was quite variable across the programmes as the process often necessitated making subjective judgments on the part of the school personnel.

Zhang (2011) opted to further explore the null effects Qiao and McNaught (2007) obtained in their analysis of the health and lifestyle behaviours. More specifically, she analysed the healthy eating, physical activity, prosocial involvement and truancy data from the RCT database using multilevel modelling statistical strategies on a much larger sample. She found that Project K participants ate breakfast and vegetables more frequently than their control counterparts at the end of the programme (statistically adjusting for baseline levels of these variables) but no differences between the groups were obtained for the other variables.
Zhang also interviewed FYD and Project K administrators and programme facilitators to assess the rationale for including improved health and fitness as a programme objective. The interviewees stated that the programme addresses health and fitness explicitly during the Wilderness Adventure by providing healthy food and organising physically challenging tasks; educational components on health and fitness are included in the Community Challenge; and all participants are required to select at least one health and fitness goal prior to the mentoring stage. The interviewees did, however, acknowledge that these behaviours may be slow to change because family members and peers have a strong influence on these. Furthermore, they reported that the primary objective of the programme was to cultivate a positive self-concept but they also felt that health and fitness behaviours could potentially be influenced indirectly through this.

Hollis, Deane, Moore & Harré (2011) used a narrative interview format to investigate the experiences of six Māori (indigenous New Zealanders) Project K graduates to elucidate the critical components of the programme and to ascertain how they felt the programme accommodated them as Māori. All of the participants credited the programme for contributing to substantial positive changes in their lives. The challenging activities that pushed their expectations, the need to work as a team, and the space created for self-reflection contributed to positive changes in their self-concept and the development of social competencies. The way instructors balanced the provision of support while promoting independence and ownership was also seen to be a positive contribution to their success as was the companionship from non-judgmental mentors.

Most of the participants noted that the programme did not accommodate them explicitly as Māori. Instead, because the same high expectations were communicated to all participants, an environment of ethnic equality was created. They appreciated this because it allowed them to temporarily escape the negative stereotype associated with being Māori. On the other hand, a couple of uncomfortable circumstances could have been avoided had the instructors been more culturally sensitive to the needs and worldviews of the young Māori participants.
Whilst the above research findings have contributed to our understanding of Project K, substantial knowledge gaps remain. Qiao and McNaught’s (2007) findings do not provide conclusive evidence about Project K’s effectiveness because of several methodological limitations (outlined in Chapter Five). The five other projects provide other important insights but the findings have not yet been synthesised. A synthesis of all Project K research findings to date is an important part of developing a “big picture” view of the programme. This thesis redresses these knowledge gaps in that it extends Qiao and McNaught’s research by conducting rigorous analyses on data collected from all of the programmes included in the RCT. It also includes the integration of findings from the five other research studies with additional qualitative data to produce a comprehensive programme theory of change. The broad aims and contributions of the thesis are also described below, following an explanation of the evaluator’s position.

**THE EVALUATOR’S POSITION**

**Research and Ontological Perspective**

I embarked on this research journey with a genuine understanding of the real-world youth development practice context, having had previous experience in youth work roles. My intention from the outset was to generate research that would be useful both to those working on the ground with young people and those in the research community. I was a stranger to Community Psychology when I began my PhD but soon learned it was a discipline that encouraged what I strived to do. That is, develop knowledge to produce positive social change while actively participating in change initiatives (Dalton, Elias & Wandersman, 2007).

Other core Community Psychology values also align with my personal beliefs of how research should be conducted. These values include promoting individual and collective wellness and social justice, demonstrating respect for human diversity, encouraging citizen participation and empowerment, collaborating with individuals, focusing on strengths, taking a socio-ecological perspective of individual development, and basing social action on empirically grounded information (Dalton, et al., 2007). Thus, I openly endorse research
approaches that attend to these values and these themes permeate the material that follows. In saying this, I aim to make this bias explicit. I also intentionally use the first person throughout this thesis as I feel this better reflects the self-reflective and active nature of the research process. It also allows for a clearer distinction between decisions that were made collaboratively with programme stakeholders and those that I take personal ownership of.

At the broadest level, I accept the post-positivist epistemology of critical realism, which has been acknowledged as falling within the extremes of the essentialist-constructionist continuum (Braun & Clarke, 2006). Critical realists propose that although a physical reality exists beyond our senses, our senses and perceptions play a role in altering this information and provide us with our own unique experiential reality alongside of the physical. Reality can be approximated by evaluating how different perceptions converge (Cook & Campbell, 1979; Trochim & Donnelly, 2008).

**Demographic and Cultural Influences**

Being that this is an evaluation of a New Zealand-based youth development programme, I necessarily comment on the New Zealand social and research context throughout the thesis. I consider how social experiences and well-being outcomes differ for individuals identifying with European, Māori, Pacific, and Asian ethnicities because these are the most prevalent ethnic groups in New Zealand society and research has demonstrated disparities in social, health, and educational outcomes between these groups (see Ministry of Social Development, 2010). With regards to these cultural distinctions, I emphasise the experiences for Māori for several reasons. First, Māori make up the second largest ethnic group in New Zealand whereas Asian and Pacific peoples form much smaller proportions. This creates practical limitations for producing accurate results specific to the latter two groups. Second, there is much more literature to draw on to explore the unique experiences of Māori than there is for young Pacific or Asian people. Third, research suggests that Māori individual and collective well-being continues to suffer because of differential exposure and access to important social determinants of health, which stems from racism that has been on-going since the colonisation of
New Zealand (Te Röpū Rangahau Hauora a Eru Pömare, 2002). Keeping with the social justice perspective that undergirds much Community Psychology research, it is important to draw attention to historical and contemporary injustices that have a pervasive impact on the health and well-being of specific groups in this country. Furthermore, New Zealand is a bicultural nation in that it is founded on an equal partnership between the British Crown and the indigenous Māori people, meaning that Māori are entitled (by way of Te Tiriti o Waitangi/The Treaty of Waitangi) to equal consideration in New Zealand research and policy issues (Te Röpū Rangahau Hauora a Eru Pömare, 2002).

Stating ethnic differences without consideration of the root causes can inadvertently reinforce attributions to individual deficits or biologically-determined factors (Te Röpū Rangahau Hauora a Eru Pömare, 2002). Thus, I would like to make clear that I am in agreement with Jones (2000, 2001) in believing that ethnic disparities result from social-environmental influences that are often linked to institutionalised, personally mediated and internalised racism, not factors inherent to the individual. I draw attention to ethnic differences because it is an important starting point from which underlying causal factors can be further explored.

In saying all of this it is important to note that my comments come from an etic (outsider) perspective as I am not a New Zealander; I am a Canadian female of French, Irish, and English descent. Although at completion of this thesis I will have lived in New Zealand for over 5 years, the lens through which I conduct my research and interpret the findings is limited by my geographic and cultural origins. The cultural information I have put forward is sourced from academic sources and is correct to the best of my knowledge. A Māori colleague from the Māori and Pacific Psychology Research Group at the University of Auckland has reviewed the sections of the thesis pertaining to Māori. I have considered the feedback and have re-framed or added some statements to incorporate her views, yet there is likely still much missing from my interpretations.
THESIS AIMS & CONTRIBUTIONS

This thesis outlines a comprehensive mixed-method evaluation of Project K. In combining a qualitative theory-driven approach (Chapter Four) with robust analyses of quantitative data collected as part of the RCT outcome evaluation (Chapter Five), I aimed to elucidate some of the potential mechanisms driving change within the programme, key moderators of programme success and the extent of the programme’s effectiveness on self-efficacy and academic achievement outcomes. My approach to the evaluation of Project K, as a whole, followed guidelines I developed as part of a novel evaluation framework (the Thoughtful Evaluation model described in Chapter Two).

The primary aim of evaluation research is to provide knowledge about a specific programme within a particular context. As such, Patton (1997) explains that the knowledge is generally not intended for broad generalisations or use beyond the localised programme scene. However, others contend that evaluation research can contribute to wider social science (Bickman, 1979; Weiss, 1998a), programming (Donaldson, 2003, 2007; Weiss, 1997a, 1998a, 1998b) and policy knowledge bases (Weiss, 1997a, 1998a, 1998b), especially when the processes, mechanisms, and conditions of change are attended to and the methods adhere to high standards of academic integrity. Chen (1990, 1994) also champions the knowledge-generation function of evaluation research that contributes specifically to improving theory and practice within the evaluation field.

This piece of research contributes to both ends. The manner in which the qualitative and quantitative methods are integrated produce scientifically rigorous results while also reflecting the authentic views of engaged stakeholders. In this way the findings reflect credible information that can be used to direct organisational decision-making, including programme and evaluation development initiatives. The findings and associated interpretations related to youth development and programming also have important implications for other individuals involved in these research and practice fields. This is particularly important for those working with youth in the New Zealand context because few rigorous New Zealand-focused studies of youth development programmes exist.
The evaluation theory and practice aspects form the second critical thread of this research story. The wide gap between evaluation theory and practice has been noted by numerous evaluation researchers (Chen, 1994; Donaldson, 2007; Shadish, Cook & Leviton, 1991; Shadish, 1998; Smith, 1993). The field has been dominated by methods-driven studies and examples of practice that are only loosely linked to evaluation theory. This has prompted calls for empirical studies investigating the relationship between evaluation theory and practice, as well as for descriptive studies outlining how evaluators integrate theory with practice and how decisions are influenced by contingencies in the practice context (Chen, 1990, 1994; Donaldson, 2007; Mark, 2003; Shadish, et al., 1991; Shadish, 1998; Smith, 1993). By first outlining the theoretical framework that guides my evaluation practice, then delineating the chronology of the evaluation process (Chapter Two), and finally reflecting on how well the theory upheld in practice (Chapter Six) I also contribute to reducing this gap. With this in mind, I move on to a description of the Thoughtful Evaluation model, the theoretical framework that underpins my approach.
Chapter Two

THOUGHTFUL EVALUATION
THE OVERARCHING FRAMEWORK

It is time to claim and name the values dimensions of our craft ~
Jennifer Greene (2011, p. 8)

EVALUATION THEORY & PRACTICE

Programme evaluations are undertaken to achieve two overarching goals: 1) Discover the truth about a programme in order to make a defensible judgment about its value and merit and 2) use that knowledge to inform decisions and future action (Patton, 1997; Shadish, 1994; Shadish, et al., 1991; Weiss, 1998a, 1998b). Yet, as explained in Chapter One, evaluations are enacted within a complex web of diverse personal and societal values; thus, what is considered “truthful” and “useful” varies substantially across evaluators and stakeholders. It stands to reason that the abundance of different evaluation approaches described in the current literature is essentially a reflection of this diversity in valuing. Therefore, before I launch into the details of the two Project K evaluation studies that form the core of this thesis, I will take the opportunity to “name and claim the values dimensions of [my] craft” (Greene, 2011, p. 8). You will see that my approach is undergirded by the values dimensions of credibility and empowerment. Together they guide a vision for my practice, which is delineated below in the Thoughtful Evaluation model. A review of the history of evaluation as a distinct field of practice and its current state of affairs helps to elucidate how this vision arose; thus, I begin here.
THE HISTORY OF EVALUATION AS A DISTINCT DISCIPLINE

Shadish and his colleagues described the history of evaluation as falling into three distinct stages: truth, use, and integration (Donaldson, 2007; Shadish, et al., 1991). Early evaluation approaches focused on “truth” through scientific inquiry, often at the expense of stakeholder “use”. In consequence, alternative approaches developed during the second stage focused much more on stakeholder inclusion and participation. Still, many critics believed the pendulum had swung too far to the other extreme, to the detriment of scientific credibility. In the current (third) stage, approaches tend to promote flexibility and situation-responsiveness. They also attempt to assimilate the strengths of the two earlier groups of approaches (Donaldson, 2003, 2007; Shadish et al., 1991). The historical progression of the field is fleshed out below.

Truth through Scientific Rigor

Evaluation as a distinct field of practice really came into its own during the 1960s. Key to the field’s proliferation was the surge of U.S. government funding for social programmes during an era known as the Great Society or the War on Poverty (Hall & Hall, 2004; House, 1993; Patton, 1997; Rossi & Freeman, 1989; Shadish et al., 1991; Weiss, 1998a). Programme funders soon realised that systematic processes were needed to determine which endeavours were legitimate and worth funding (Hall & Hall, 2004; Patton, 1997, Shadish et al., 1991; Weiss, 1998a). With social and behavioural scientists realising that research methods could be applied to gain knowledge of real-world problems, the solution became apparent. These methods could be utilised to determine the worth and merit of the different programmes (House, 1993; Patton, 1997; Rossi & Freeman, 1989).

Because the interest in evaluation during this early stage was focused on discerning the success or failure of government-funded programs, evaluation questions were centred on whether programme objectives were being met (i.e. outcome evaluation), how programmes fared in comparison to each other (comparative assessments), and which were cost-effective enterprises (cost-benefit analyses). In essence, decision-makers wanted judgments that “summed up” the
overall success of an initiative. This led to what is now known as summative evaluation (Hall & Hall, 2004; Patton, 1997, 2002).

Social science at this time was steeped in a positivist single-reality worldview borrowed from the natural sciences, and methodologies primarily aligned with the quantitative-experimental paradigm (Hall & Hall, 2004; House, 1993; Mark, 2003; Patton, 1997, 2002). According to this view, the only way to obtain veritable knowledge is through the measurement and manipulation of variables, and direct, objective observation of cause and effect (Trochim & Donnelly, 2008). Essentially, “truth” about a programme’s success was generally only seen to be valid if the knowledge was yielded from the scientific method and experimental designs (Patton, 1997, 2002).

Randomised controlled trials, as a particular type of experimental design, were (and, according to many, still are) seen to be the gold standard of evaluation designs because of the superiority they afforded for making causal claims about a programme’s effects (Campbell & Stanley, 1963; Cook, 2000; Cook & Campbell, 1979; Lipsey & Cordray, 2000; Rossi & Freeman, 1989; Shadish, Cook, & Campbell, 2002; Trochim & Donnelly, 2008; D. Stufflebeam, 2001; Weiss, 1998a). Nevertheless, evaluators who were steadfast in their adherence to the quantitative-experimental paradigm were faced with a major backlash from programme stakeholders and other evaluators who criticised the approach on numerous grounds (Hall & Hall, 2004; House, 1993; Patton, 1997; Shadish et al, 1991; Weiss & Rein, 1969, 1970).

First, the great majority of programmes evaluated in this manner were found to have no effect (House, 1993). However, programme practitioners and some evaluators reported a different story when observing some of these programmes in action (Patton, 1997, 2002). In essence, the “truth” about complex programmes was being reduced to performance on a small number of measurable performance indicators that failed to capture the complex nature of programmes (Patton, 1997) and the differing views of stakeholders (House, 1993; Patton, 1997, 2002). This problem was exacerbated by evaluators who distanced themselves from the stakeholders and programme operations (and thus a genuine understanding of
the programme) in order to maintain their objective stance. Information about the programme was communicated in scientific and statistical terms thus reports were often incomprehensible and/or irrelevant to stakeholders, further alienating them from the evaluation (Patton, 1997).

What is more, evaluators realised that implementing social experiments within real-world settings was much more difficult than in a research lab (House, 1993; Hall & Hall, 2004; Patton, 1997; D. Stufflebeam, 2001; Weiss & Rein, 1969, 1970). Reported difficulties included sabotage of the implementation process by uncooperative practitioners (Weiss, 1998a), participants receiving differential exposure to programme services, differential attrition of programme and control participants over the course of the evaluation and control participants receiving similar services (Lipsey & Cordray, 2000; Weiss, 1998a). In the end, randomisation process was often tainted (see Weiss & Rein, 1969, 1970 for an example).

Additionally, evaluations that focused exclusively on programme effectiveness generated no insights about how or why a programme did or did not work (Chen, 1990; Patton, 1997; D. Stufflebeam, 2001; Weiss & Rein, 1969, 1970). This created difficulties for those wanting to replicate the successful aspects of a programme or for those wanting to make improvements. Accordingly, they came to be known as “black box” approaches because people could not discern what was occurring within the programme to produce or negate change (Astbury & Leeuw, 2010; Campbell, 1970; Chen, 1990; Donaldson, 2003, 2007). These factors all contributed to the evaluation “utilization crisis” (Patton, 1997, p. 7) that soon followed the initial growth in programme evaluation. Evaluations were simply not being used.

**Use through Stakeholder Involvement and Attention to Processes**

The “utilization crisis” prompted a series of alternative approaches based on attention to programme processes and involving stakeholders in the evaluation process (Hall & Hall, 2004; House, 1993; Mark, 2003; Patton, 1997; Weiss, 1998a). Process-focused approaches undertaken to provide formative feedback about the effectiveness of different components of programme design, implementation, and
delivery were more informative in terms of programme improvement and expansion. They therefore became legitimate forms of evaluation (Hall & Hall, 2004; House, 1993; Patton, 1997; Weiss, 1998a). A notable shift in researcher-stakeholder relations also eventuated as researchers began to recognize the importance of respecting and listening to those directly involved with programme initiatives. This increased the likelihood of programme designers and practitioners picking up recommendations (Dalton et al., 2007; Fetterman, 2003; Fetterman & Wandersman, 2005; O’Leary, 2005, Patton, 1997; Weiss, 1998a).

A variety of participatory approaches followed suit and Weiss (1998a) later classified these along a participatory continuum. Stakeholder evaluations are the most conservative of these. Decisions about which evaluation questions to pursue and which methods are suitable are decided collaboratively, but the evaluator retains control over implementing the evaluation and analysing the data. In collaborative evaluation stakeholders are co-investigators in the evaluation and may take part in data collection, analysis and interpretation. Empowerment evaluation is the most liberal (and contentious) of the participatory approaches as stakeholders take over all aspects of the evaluation (Weiss, 1998a) and the evaluator remains on the side-line as a “critical friend” and facilitator (Fetterman, 2003; Fetterman & Wandersman, 2005).

Related to the above progress of process-focused and participatory approaches, the qualitative-naturalistic paradigm began to gain some ground (House, 1993; Patton, 1997, 2002; Weiss, 1998a). Proponents of this methodological camp highlighted the advantages of qualitative methods to provide a holistic and in-depth picture of the programme experience, to identify outcomes not easily captured with standardised quantitative methods (Patton, 1997, 2002), and to involve stakeholders and infuse their voices into the evaluation (House, 1993; Patton, 1997, 2002; Weiss, 1998a).

Nonetheless, these alternative approaches were far from being unanimously accepted. The more liberal participatory approaches (i.e. empowerment evaluation) were seen to lack rigor and credibility as they were thought to reflect
the self-interested biases of the stakeholders who were involved (see Stufflebeam, 1994). There was also a visible clash of heads between staunch advocates of the quantitative-experimental paradigm and those favouring a qualitative-naturalistic framework, launching the long-standing “paradigm wars” (Weiss, 1998a, p. 14; see also House, 1993; Mark, 2003; Patton, 1997, 2002).

**Methodological Integration and Contingency-Based Approaches**

Though a few loyal adherents may still plant their feet firmly in a particular paradigmatic camp, the debate has come a long way (House, 1993; Mark, 2003; Patton, 1997, 2002; Weiss, 1998a). There has been shift towards accepting that phenomena that are difficult to measure and observe directly are still worth investigating (Trochim & Donnelly, 2008) and that evaluator subjectivity is inevitable. Rather than distancing themselves from the evaluand, evaluators are now encouraged to lay their philosophical and values perspective on the table from the outset and to try to maintain a neutral stance (Patton, 1997, 2002; Trochim & Donnelly, 2008). All in all, there is now a general consensus that both formative/process-focused and summative/outcomes-focused evaluation and qualitative-naturalistic and quantitative-experimental approaches can serve different but complementary functions (House, 1993; Patton, 1997, 2002; Weiss, 1998a).

More recent models of evaluation tend to endorse a pragmatic or flexible approach that draws on aspects of both quantitative and qualitative paradigms depending on the suitability of each to the purpose and the context of the evaluation (Chen, 1994; Donaldson, 2003, 2007; Mark, 2003; Shadish et al., 1991). Shadish et al. (1991) appropriately called these contingency-based evaluation theories. Stakeholder consultation and, at least peripheral involvement, has become the standard (House, 1993; Mark, 2003; Patton, 1997, 2002; Weiss, 1998a). Furthermore, whether quantitative or qualitative, a systematic approach and high levels of competency in the selected methods are now universal expectations of the field. That said, debate over the role of the evaluator still exists (Shadish, et al., 1991; Stufflebeam, 1994; Weiss, 1998a). At the turn of the century many theorists maintained that the high degrees of stakeholder involvement seen in some of the
evaluation capacity-building approaches should be tempered if the evaluator’s role is to render a judgement about the programme’s merit and worth (Chen, 1994; Shadish, 1994; D. Stufflebeam, 2001; Stufflebeam, 1994; Weiss, 1998a).

THE ESSENTIAL FEATURES

To sum what is discussed above, views about evaluation practice are wide-ranging and this has been the cause of substantial friction between evaluators in the past. Tensions still exist but substantial progress past the paradigms debate has been made. Many evaluators across the globe have come to some common ground regarding the essential features of high quality evaluations.

To encourage high quality evaluation practice and greater professionalisation of the field, some of these commonalities have been explicitly articulated as basic principles and practice guidelines by several evaluation organisations. Two of these, the Program Evaluation Standards developed by the Joint Committee for Standards on Educational Evaluation (JCSEE, 2011) and the American Evaluation Association’s (AEA) Guiding Principles (AEA, 2004), have had widespread influence and serve as the basis for the development of frameworks in other regions of the globe, for instance the Australasian Evaluation Society’s (AES) Guidelines for Ethical Conduct of Evaluation (AES, 2010) and the recently developed Aotearoa New Zealand Evaluation Association’s (ANZEA) Framework for Evaluator Competencies (ANZEA, 2011). A synopsis of the Program Standards and the Guiding Principles are provided below followed by a brief overview of how evaluation should be practiced in Aotearoa New Zealand according to ANZEA’s competencies framework.

Joint Committee Standards for Educational Evaluation (JCSEE)’s Program Evaluation Standards

In 1981, the Joint Committee on Standards for Educational Evaluation developed a list of recommendations required to conduct a sound evaluation (these were revised in 1994 and again in 2011; www.jcsee.org). Originally devised as guidelines for the evaluation of educational programmes specific to the USA and Canada (Stufflebeam, 1994), these standards have now become commonly
accepted across the entire program evaluation field (Patton, 1997; Stufflebeam, 1994). Thirty standards fit within five categories of utility, feasibility, propriety, accuracy and accountability (the latter standard was added in the most recent revision; see Yarbrough et al., 2011).

**American Evaluation Association (AEA)’s Guiding Principles**

The five inter-related guiding principles of systematic inquiry, competence, integrity/honesty, respect for people, and responsibilities to general and public welfare were developed by an AEA Task force and officially endorsed by the association in 1994. They were most recently revised in 2004. It is acknowledged that these were developed within a western and, more specifically, U.S. context; thus, they may be limited in their application in other cultures ([www.eval.org](http://www.eval.org)).

**Aotearoa New Zealand Evaluation Association (ANZEA)’s Evaluator Competencies**

The AEA’s Guiding Principles and the JCSEE’s Program Evaluation Standards form core assumptions in ANZEA’s (2011) recently developed Evaluator Competencies framework. However, because Aotearoa New Zealand is unique in its socio-political and cultural landscape, the principles of quality practice differ from these guidelines in the central emphasis placed on cross cultural competence and the commitment to the *Te Tiriti o Waitangi* /The Treaty of Waitangi (more on this in Chapter Three). Specifically, competent evaluation practice in this region must reflect the tripartite principles of partnership, protection, and participation that underpin the bicultural agreement between the indigenous Māori and Pākehā (non-Māori New Zealanders, generally of European descent). The framework consists of four overlapping core domains: contextual analysis and engagement; systematic evaluative inquiry; evaluation project management and professional evaluation practice; and reflective practice and professional development (ANZEA, 2011).

Table 1 outlines the features of each framework. In reviewing the groupings, it is clear that the JCSEE, AEA and ANZEA endorse approaches that incorporate systematic and technically adequate methods; stakeholder consultation and consideration; ethical, respectful, and reflective practice; and detailed, transparent,
and understandable communications to produce credible and defensible findings and useful social contributions.
Table 1. The essential features associated with high quality evaluation practice

<table>
<thead>
<tr>
<th>Program Evaluation Standards</th>
<th>Guiding Principles for Evaluators</th>
<th>Evaluator Competencies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy:</strong></td>
<td><strong>Systematic Inquiry:</strong></td>
<td><strong>Systematic Evaluative Inquiry:</strong></td>
</tr>
<tr>
<td>Decisions and conclusions are explicitly justified; interpretations are valid; information is dependable and consistent; documentation has sufficient detail of programme and context; technically adequate designs and analytic strategies are employed; evaluative reasoning is explicit; scope of communications is adequate and biases, errors, misconceptions are guarded against</td>
<td>Highest technical standards are adhered to; strengths and opportunities for improvement are explored with client; communications include sufficient details of approach, methods, and limitations to all critique</td>
<td>Evaluator is skilled and knowledgeable in evaluation design, systematic data collection, critical thinking, analysis and synthesis; interpretations are valid and defensible; methods, conclusions and judgments are transparent; findings are reported in credible, useful and actionable way;</td>
</tr>
<tr>
<td><strong>Feasibility:</strong></td>
<td><strong>Competence:</strong></td>
<td><strong>Evaluation Project Management and Professional Practice:</strong></td>
</tr>
<tr>
<td>Effective project management; practical and responsive procedures; diverse cultural and political interests recognised and attended to; effective and efficient use of resources</td>
<td>Evaluation team collectively has appropriate level of knowledge, skills and experience; cultural competence is demonstrated; activities are conducted within limitations of competence; limitations of knowledge/skills acknowledged; knowledge and skill development sought when needed</td>
<td>Evaluator is able to effectively manage the project; to build collaborative and respectful relationships with stakeholders and other team members; appropriate standards and ethics of professional evaluation practice are adhered to</td>
</tr>
<tr>
<td><strong>Accountability:</strong></td>
<td><strong>Integrity/Honesty:</strong></td>
<td><strong>Reflective practice and professional development:</strong></td>
</tr>
<tr>
<td>Evaluations are fully documented; processes and outputs are measured against standards; and external meta-evaluation is encouraged</td>
<td>Costs, methods, limitations, scope of results are negotiated in an honest manner; conflicts of interest are disclosed; changes to proposed plan are recorded and reported with implications; interests and values are made explicit; procedures, data, findings are accurately reported and misuse prevented; concerns about misleading information are resolved; financial sources of support and source of request for evaluation are made clear.</td>
<td>Evaluator reflects on own areas of expertise, identity and approach to practice; personal growth areas are identified; evaluator engages in professional development opportunities; and contributes to the profession</td>
</tr>
</tbody>
</table>
Thoughtful Evaluation

Guidelines for Conducting High Quality Evaluations

**Program Evaluation Standards**  
Joint Committee on Standards for Educational Evaluation (2011)

**Guiding Principles for Evaluators**  

**Evaluator Competencies**  
Aotearoa New Zealand Evaluation Association (2011)

<table>
<thead>
<tr>
<th><strong>Propriety:</strong></th>
<th><strong>Respect for People:</strong></th>
<th><strong>Contextual analysis and Engagement:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsive and inclusive of varied stakeholder groups; formal agreements are negotiated and adhered to; human and legal rights are protected; information is understandable and fair; full disclosure and transparency about process and findings; conflicts of interest are declared; activities are conducted with fiscal responsibility</td>
<td>Understanding of broader context is sought; ethical standards are adhered to; benefits are maximised and any potential harms considered and reduced; stakeholder dignity is respected during process and in communications; bidirectional rewards between all those involved are encouraged where possible; diverse stakeholder characteristics and values are considered and respected</td>
<td>Wider context and evaluation situation is identified, understood, articulated, and considered; evaluation team has credibility in the evaluative context and builds respectful relationships with stakeholders; evaluators engage with relevant individuals in wider context when conducting process so interpretations are credible and valid to all affected stakeholders</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Utility:</strong></th>
<th><strong>Responsibilities to General and Public Welfare:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluators are credible and qualified; full range of stakeholders attended to; purposes are negotiated and stakeholder needs addressed; explicit articulation of values; information is relevant; processes and products are meaningful; dissemination is timely and appropriate to audiences; concern demonstrated for consequences and influence of evaluation</td>
<td>Full range of stakeholder perspectives and interests attended to; broad assumptions, implications, and side effects are considered; stakeholders are given access to information and this is actively disseminated in an understandable form; client and stakeholder needs are balanced; interests and welfare of society as a whole are considered</td>
</tr>
</tbody>
</table>
It appears that, at least historically, programme evaluators have struggled to reconcile a gap between the values of the scientific or research community and the values of those closest to the programmes being evaluated. This struggle has been repeatedly acknowledged to stem from a critical tension between evaluation credibility and evaluation use (Hall & Hall, 2004; Patton, 1997; Rossi & Freeman, 1989; Shadish, 1994; Shadish et al., 1991; D. Stufflebeam, 2001; Stufflebeam, 1994). Here, I argue that the tension arises from the difficulty of conducting evaluations in a manner that yields scientifically credible information while at the same time empowering programme stakeholders. I believe that an approach that focuses on this dual emphasis will generate greater evaluation use.

Evaluation use has been a prominent issue in evaluation theory since the “utilisation crisis” in the early stages of its history. Initially, discussions centred on instrumental use which involves using the evaluation findings to make decisions about the programme. Conceptual use was subsequently recognised as an important contribution even if evaluation recommendations were not immediately enacted. This refers to the new understanding of concepts and programme evaluation issues that can arise as a result of the evaluation for people directly involved with the programme (Patton, 1997; Weiss, 1998b). Later, Patton (1997) broadened the conversation to include process use, which referred to the learning that resulted from actually being involved in the evaluation process.

Weiss argues (1998a, 1998b, 1997a) that social science researchers, policymakers, evaluators and practitioners involved in other similar programmes are legitimate audiences to consider when thinking about an evaluation’s use. Information that contributes to the wider fields of social programming and evaluation policy and practice can generate widespread use through “enlightenment” (Weiss, 1998b, pg. 24). Several advocates of the theory-driven perspective join Weiss in promoting the knowledge-generation function of evaluations for developing “cumulative wisdom” (Donaldson, 2003, p. 134) in the social sciences (Chen, 1990; Chen & Rossi, 1980, 1983, 1987; Donaldson, 2007). It
follows that concepts of use should not be constrained to that which is specifically relevant to programme stakeholders. Yet, the topic has maintained a general focus on use by primary stakeholders (most often referring to program designers, practitioners and funders).

Below, I present my Thoughtful Evaluation model. As stated, it is a model that emphasises the balance between scientific credibility and stakeholder empowerment. It draws strongly on approaches promoted by Donaldson (2003, 2007), Patton (1997) and Weiss (1997a, 1997b, 1998a, 2000) but is also informed by the historical debates of the evaluation field, the JCSEE’s Program Evaluation Standards and the AEA’s Guiding Principles. I believe it also aligns well with the four domains of ANZEA’s Competency framework, despite these being published after my development of the model.

THE THOUGHTFUL EVALUATION MODEL

Stake (2004) described his model of Responsive Evaluation as “an attitude rather than a model or recipe” (p. 86) and the same notion applies to Thoughtful Evaluation. It is a visual representation (see Figure 1) that serves as a tool for self-reflective meta-evaluation, or the evaluation of one’s own evaluation. For this reason, a thoughtful evaluator is positioned in the centre of the model, taking responsibility for balancing the two dimensions of credibility and empowerment.

The two dimensions of the Thoughtful Evaluation model emerged after analysing the constituent elements of the JCSEE’s Program Evaluations Standards and the AEA’s Guiding Principles. I found that these elements could be divided into themes of scientific credibility or stakeholder empowerment, though there was evident overlap between the dimensions. These dimensions are embedded within a larger concept of thoughtfulness (on the part of the evaluation researcher), referring both to a critical, reflective and analytical approach and genuine consideration of others’ needs and values. It is the position of the evaluator, as the conscientious puppeteer that carefully calibrates the balance between the credible and empowering aspects that distinguishes this approach. I return to this after a description of the two dimensions.
Figure 1. The Thoughtful Evaluation Model illustrating the role of the evaluator in balancing the credible and empowering aspects of evaluations to meet the needs of external and internal programme stakeholders.

Model Dimensions

The Credibility Dimension

Features within this dimension align with ideas of research integrity and credibility outlined in the basic and applied research methods literature (e.g. Trochim & Donelly, 2008). This includes the importance of systematic data collection, rigorous analysis and appropriate interpretation, including balanced reporting of strengths and weaknesses to yield results that are accurate and valid. These aspects are encompassed by the broader ideals of technical adequacy and in some instances may contribute to generalisable and reproducible findings. Information acquired in this manner may also contribute to broader theories and wider practice contexts if interpreted in a manner that shows extensions beyond the localised programme context. In view of this, and because these features arise from the research methods literature they tend to target the needs and values of those external to the programme.

The Empowering Dimension

Empowerment is a central tenet of community psychology (Dalton et al., 2007). When referring to individuals, psychological empowerment occurs as a result of being involved in and taking advantage of empowering processes. Empowering processes are those which: increase the knowledge one has about his
or her environment; develop skills so that one can make critical decisions this environment; and encourage active participation in decisions that will affect oneself in this environment. This concept can be applied to any situation where one seeks to close the gap between the current situation and the ideal (Zimmerman, 2000).

Empowerment is also a concept described in the organisational and management research literature. In this context, empowering processes are described as those in which leadership and decision-making are shared across different levels of the organisation, where employees feel meaningfully connected in their responsibilities, and the workplace is seen as a flexible and dynamic environment where any individual could have an impact (see Arnold, Arad, Rhoades and Drasgow, 2000; Konczak, Stelly and Trusty, 2000; Matthews, Diaz and Cole, 2002). Whether based in community psychology or business initiatives, the outcome of an empowering environment is increased participation in meaningful activities to create positive change towards individual or organisational self-determination.

The parallels between this idea of empowerment and the idea of stakeholder participation and use described in the evaluation literature are clear-cut. Use is said to increase if there is mutual respect between evaluators and stakeholders for their particular areas of expertise, if decisions are shared, and if the approach is inclusive of the range of stakeholder groups. In this way the evaluator engages with the values and needs of diverse stakeholder groups to ensure that the decisions and recommendations are culturally appropriate and relevant. These are all essentially issues of empowerment because they cultivate stakeholder knowledge of the programme environment and encourage active participation in processes that promote positive organisational change.

**Dimensional Balance, Tension and Overlap**

The balance of credibility and empowerment may differ according to influences (both personal and situational) and the specific research purpose (which will evidently be linked to the values of the individuals/organisation that
commissioned the evaluation). To illustrate, an outcome evaluation done by an external evaluator may be viewed as high in credibility but low in empowerment if stakeholder involvement is limited to brief consultation prior to evaluation implementation and dissemination of the findings at the completion of the process. In contrast, an empowerment evaluation that allows stakeholders to control the full evaluation process will obviously be viewed as being high in empowerment but may be seen to be low in credibility if stakeholders lack a true understanding of the validity issues associated with different research methods.

Attention to particular aspects of scientific credibility may also detract from the empowering aspects and vice versa because of the inherent tension between some features of the two. For instance, experimental designs are lauded for their methodological rigor; yet, they can take a substantial amount of time from the design to the dissemination stage and timeliness is essential to producing relevant results for stakeholders (Weiss, 1998a; Patton, 1997). The long wait time for results that may no longer be applicable at later stages of the programme can potentially engender feelings of disempowerment for stakeholders. Likewise, the intensive nature of the data collection process can interfere with programme practitioners’ other responsibilities, especially if they are involved in with data collection efforts (Weiss, 1998a). This can reduce any involved stakeholders’ sense of agency in being able to effect positive change and, as a corollary, their cooperation with the process.

Nonetheless, there are also times when attention to particular aspects of one dimension enhances aspects of the other. The shared space at the centre of the model represents these overlapping mutually-enhancing features. An example of this is the triangulation of multiple stakeholder perspectives. This strengthens the scientific credibility and validity of qualitative findings (Patton, 2002) but also necessarily entails involving stakeholders in the research process and giving voice to their respected opinions. Focusing on some of these mutually-enhancing features to maximise the benefits of each dimension is a key part of a Thoughtful approach.
Holding the Balance: The Evaluator’s Imperative

As touched on above, at the beginning of the 21st century evaluators were still questioning (Chen, 1994; Shadish et al., 1991) or promoting (Stufflebeam, 1994; Weiss, 1998a) the separation of evaluator roles according to what could be seen to be dimensions of scientific credibility and stakeholder empowerment. Nevertheless, the JCSEE Program Evaluation Standards and the AEA Guiding Principles clearly endorse attention to what I consider to be credible and empowering features for all professional evaluations. In the most integrative contemporary approaches stakeholders are consulted or involved at different parts of the evaluation process and rigorous methods are promoted (Donaldson, 2003, 2007; Mark, 2003; Shadish et al., 1991). The ideas of stakeholder empowerment and methodological rigor have also been explicitly touched on by several evaluation theorists in the recent past, and these individuals allude to a need for balance between the two (Donaldson, 2003, 2007; Fetterman, 2003, Fetterman & Wandersman, 2005; Hall & Hall, 2004; Patton, 1997; D. Stufflebeam, 2001). Overall, however, the essentiality of the balance is not overtly advanced in the current literature, nor is the need for the evaluator to take ownership of that role.

Moreover, today it is still not unusual to see the separation of methodological roles within a single summative evaluation, such that the individual involved in the analysis may not necessarily be involved in the design, data collection process or the dissemination of the results. The MSD-funded RCT evaluation of Project K is a prime example. Qiao and McNaught (2007) conducted analyses on the data and produced the report of the findings but were not involved in the design of the evaluation, the measures, or the data collection process and thus were far removed from the programme scene. This piece-meal approach may inadvertently diffuse the responsibility for any one person to attend to both the credibility and empowering features. Careful reflection on the appropriateness of the analytic strategy and on the wider implications of reporting the results in a particular manner is as critical as reflection on the diverse values of different stakeholder groups when deciding how to focus evaluation questions. Likewise, a thorough understanding of the credibility of the methods used is as essential to
appropriate dissemination of the evaluation information as is employing the methods to begin with. This is the crux of Thoughtful Evaluation.

The Thoughtful Evaluation model was developed with single-evaluator situations in mind, as this was my situation when I began the research process. The model may, therefore, not be ideal in large-scale evaluations where evaluative activities must be divided amongst individuals working on an evaluation team. However, the lessons of the past have taught us that use is maximised when both dimensions are more or less equally attended to. Hence, I believe that, regardless of the nature of the evaluation, one evaluator should take responsibility for ensuring this is some degree of balance between scientific credibility and stakeholder empowerment across the different components of an evaluation. By following the Thoughtful Evaluation approach one is encouraged to serve the evaluation needs of various audiences both internal and external to the programme, thereby increasing the evaluation’s potential use. As stated, the Thoughtful Evaluation model provides the vision for the current evaluation of Project K.

**CHRONOLOGY OF THE EVALUATION PROCESS**

The two core studies of this thesis are presented in Chapter Four and Five. The first of these presents an attempt to construct a “white box” (Astbury and Leeuw, 2010, p. 364) view of the programme through the development of Project K’s theory of change. The second takes a more traditional “black box” perspective to evaluate Project K’s effectiveness on self-efficacy and academic achievement outcomes using data collected as part of a randomised controlled trial. The former study is particularly strong in empowering characteristics while the credibility aspects feature strongly in the latter, yet I endeavoured to balance both dimensions within each study.

It would be misrepresentative, however, to imply that the thoughtful approach was restricted to two discrete studies. Attention to scientific credibility and stakeholder empowerment characterised the overall evaluation process and permeated aspects of the research process beyond what is presented in Chapters
Four and Five. In addition, the presentation of the studies in this order does not accurately reflect how the process actually unfolded. As stated in Chapter One, the RCT was well underway when I came on board. For this reason the chronological development of the research process is presented in diagrammatic form below followed by a description of the evaluation process prior to moving into Chapter Three, where an in-depth contextual analysis of the youth development and programming field overseas and in New Zealand is presented.
Figure 2. The chronology of the research process illustrating the concurrence of the two core thesis studies, the order of the various literature reviews, and the on-going nature of the stakeholder consultation and evaluator self-reflection processes.
My role was not defined when FYD agreed to my involvement in Project K’s evaluation activities. There was a general understanding that I would be involved in investigating some of the RCT data but, as it was an extensive dataset, the nature of the research questions had not been determined. After reviewing the evaluation literature and developing my evaluation framework, I approached the organisation with my *Thoughtful Evaluation* model and indicated that I was interested in working collaboratively with FYD to decide the focus of the evaluation. I also proposed that we take a strong participatory approach by extending an invitation to other programme stakeholder groups to become actively involved in the research process.

When I made the suggestion to a group of Project K staff working in different regions of the country at an official *hui* (gathering or meeting) in early 2008, I encountered resistance to any notion of further evaluation. Their extensive involvement in the data collection for the RCT data until that point had substantially increased their workloads and reduced their available time for other programme activities. Understandably this had generated frustration with programme evaluation in general. I did, however, get a sense that my involvement in the analysis of the RCT data would be acceptable, as long as this did not require much additional participation on their part.

I recognised that if stakeholder involvement was not volitional, the process would end up being disempowering. However, I still hoped some stakeholders would consider being involved in determining the focus of my particular research questions and in other more peripheral ways to ensure the questions were relevant and the findings would be accurate, representative of multiple perspectives, and useful. Thus I reconsidered how they could be included without creating an additional burden on their time and energy.

Not long after, I travelled to the majority of the Project K regions to meet with Programme Directors (and any other interested staff members) and inquired further into their apprehension to evaluation. Several individuals expressed concerns with particular aspects of the RCT data collection process and some were
not convinced that the findings would be relevant to their particular programme. I agreed to review these issues, to assess alternatives and to produce reports of the findings specific to each Project K region as well as for the national programme as a whole. FYD later agreed to all proposed changes to the RCT evaluation process and regional reports were produced.

During these meetings, I also proposed two initiatives that would allow their peripheral involvement in the evaluation process and a part in the decision-making process. First, I offered the development of a virtual communication network, where interested stakeholders could determine their own level of participation at any stage of the process. They would be able to receive information about the evaluation as reports were developed and would have direct access to me if they had questions or concerns and they would be encouraged to provide feedback.

Second, I presented Programme Directors with an evaluation questions voting survey. The survey questions were uplifted from D. Stufflebeam’s (2001) article describing 22 different approaches to evaluation. The questions were modified to suit the Project K context and were categorised by theme and purpose to form 30 different voting options (see Voting Survey in Appendix F). Programme Directors were informed that the purpose of the exercise was to assess their information needs and that these (along with considerations of project feasibility) would inform my research direction, including which data from the RCT sample to focus on and other research avenues which could help to uncover the desired information. They were also encouraged to distribute the survey to other stakeholders.

In total, 67 stakeholders (including FYD staff, Project K staff, Trustees, Project K participant caregivers, mentors, and other volunteers) took part in the voting process. The voting results pointed to a desire for more information about programme processes (e.g. How and why does the programme produce outcomes?; What changes would produce better outcomes? Is it reaching the right students?) and about effectiveness (Is Project K working? To what extent were
Project K objectives achieved?), especially for different types of participants (For which types of young people does Project K work best and why?).

I determined that the development of a logic model detailing the programme’s theory of change would be the best place to start in terms of ascertaining the key programme processes and potential moderating factors (including participant characteristics). I also took the votes relating to programme effectiveness as confirmation that I should investigate the outcome data collected as part of the RCT and where possible to determine if it was differentially effective for different participant subgroups. I produced a report for the stakeholders involved detailing the results of the voting process and a proposal of how I would approach the questions. Stakeholders were encouraged to ask questions and to submit feedback on the proposal.

I negotiated the terms of my involvement in Project K’s evaluation with FYD (see the Memorandum of Understanding and Organisational Consent Form in Appendix A) and the data analysis (and the necessary up-skilling in the appropriate statistical methods) occurred concurrently with the development of the logic model. The logic model building process included a workshop with FYD National Support Office Staff and travel to almost every Project K region to conduct focus groups with Project K staff members. Youth participants were not directly involved but their qualitative comments from previous evaluation surveys were analysed (more on this in Chapter Four).

Preliminary results from the RCT data analysis study were presented via written reports to FYD and to the Project K regions (these included specific regional and National programme reports). Results were also presented to National Support Office staff at an FYD research meeting. The preliminary logic model was presented to interested National Support Office and Project K staff at another FYD hui (during the latter part of 2008) and their feedback was solicited at this time. All presentations to external groups (e.g. academic conferences) based on the results of these studies were shared with and reviewed by FYD prior to wider dissemination. Furthermore, the key findings from both studies were presented to
National Support Office Staff, Project K Programme Directors and contracted providers of the Wilderness and Community components at a hui in 2011. At this time, written feedback from individual stakeholders on each component of the logic model was solicited. Programme development opportunities identified through the research process were also discussed and a brainstorming workshop activity was facilitated to get stakeholders involved in thinking through action-oriented strategies for programme improvement. The suggestions from stakeholders were compiled (see Table 1 in Appendix F) and included as part of a summary report of the feedback process to FYD.

The logic model findings also pointed to potentially important programme process, moderator, and outcome variables that were not being measured (e.g. autonomy support from facilitators and mentors, participant engagement, and sense of community). As a result, FYD’s Evaluation Coordinator and I collaboratively sourced relevant and psychometrically sound measures of several of these constructs; adapted them to suit the Project K context; and revised the Project K evaluation process for programmes operating from 2010 to focus on testing these mechanisms of change.

The above chronology simply delineates the evaluation process at a descriptive level to provide the reader with a better understanding of the key events and decisions that led to the research projects outlined in Chapter Four and Five. A review of the relevant youth development programme literature is presented in Chapter Three to situate the project within the broader research context for the reader but, as shown in Figure 2, the literature review occurred after the development of the logic model described in Chapter Four because it served to theoretically validate the programme theory (more on this in Chapter Four). In Chapter Six, I come full circle and draw on the stakeholder engagement processes described above processes to evaluate how well my approach met the aims of the Thoughtful Evaluation model. The final chapter (Seven) concludes with a grand discussion that highlights the strengths of the research, the implications of the lessons learned and the directions for future research and practice.
Chapter Three

POSITIVE YOUTH DEVELOPMENT
THROUGH ADVENTURE, SERVICE & MENTORING

FYD’s programmes help young people discover possibility – creating a youth population with a positive outlook and eyes open to the future ~ The Foundation for Youth Development (www.fyd.org.nz)

ADOLESCENT DEVELOPMENT & YOUTH PROGRAMMING

That adolescence is a time of great change is uncontested. During this life phase, individuals have to contend with major changes across multiple domains of identity. The more salient of these are the biological transformations that take place but individuals must also negotiate cognitive, emotional, and socio-cultural changes during this time (McLaren, 2002; Steinberg, 2008). Due to these often concurrent transformations, adolescence may be considered the biggest life transition of all and this has attracted the attention of researchers eager to utilise this complex period of growth as a “natural ontogenetic laboratory” (Lerner, 2005, p.10) for understanding human life-span development in general.

The road to adulthood involves a series of key developmental tasks. As adolescents become more autonomous and shift from spending a great deal of time with family to more time with peers, the development of interpersonal skills becomes critical. The need for autonomy, however, must be balanced with the maintenance of positive family ties as families continue to provide vital resources for positive growth. Coping with the physical changes associated with puberty also necessitates coming to terms with a sexual identity and often, assuming a role in a romantic relationship. The development of abstract thought and metacognition create expectations of self-regulation and decision-making skills that play a prominent role in negotiating safe sexual behaviour, avoiding risks, academic
achievement and moral development. The increasing scholastic demands that occur with major school transitions, as well as the exploration and eventual selection of a vocational path are other challenges that must be contended with on one’s journey to an integrated personal identity (McLaren, 2002; Steinberg, 2008).

Bumps in the road are inevitable as one navigates through these numerous developmental markers but for the majority of adolescents it is not the period of “storm and stress” that G. Stanley Hall proclaimed when the first research on adolescence was put to paper in 1904 (Arnett 1999; Lerner, 2002). Nevertheless, the focal point of the first stage in the history of adolescent research was on remedying pathological populations of adolescents (Catalano, Berglund, Ryan, Lonczak & Hawkins, 2004; Damon, 2004; Lerner, 2005) and research focused solely on fixing problems is still prolific. Consequently, this deficit view fuels the negative portrayal of youth in the media and, subsequently, in the public mind-set (Benson, 1997; Crane, 1994; Damon, 2004; Giroux, 1996; Mahoney & Lafforty, 2003). In contest to this inclination, a new approach began to be advocated in the 1990’s – the Positive Youth Development (PYD) approach (Catalano, et al., 2004). Acknowledging that a holistic understanding of adolescent development must include the adverse aspects, the PYD approach takes the perspective that all young people have inherent strengths (Damon, 2004) and that a healthy adolescent does not equate to one that is simply without problems (Lerner, 2005).

A POSITIVE AND ECOLOGICAL APPROACH TO DEVELOPMENT

Until recently, another trend in psychological research was a focus on understanding individuals as isolated from their context (Lerner, 2002; Lerner, Fisher, & Weinberg, 2000) whereas proponents of the developmental systems approach argue that human development occurs through a series of dynamic interactions between an individual and the multiple levels of context in which he or she is embedded. These interactions are reciprocal: humans are active in influencing their environment and the multi-layered ecological system of which they are a part shapes their functioning and growth (Lerner, 2002; Lerner,
Dowling, & Anderson, 2003); thus, research on human development should account for these person-context relationships (Lerner, 2002).

Urie Brofenbrenner’s (1979) seminal work on the ecology of human development distinguished these multiple layers according to four levels differing in proximity to the individual. The most proximal, the microsystem or the immediate setting of the individual, for adolescents would largely be accounted for by the home, school and community. The mesosystem signifies the interrelations between the settings in the microsystem and the exosystem are the influences on the individual that arise from interactions between contexts do not directly involve the individual but that influence the immediate contexts of which they are a part (e.g. a parent’s work environment has an indirect influence on a young person’s home life). The final level is the macrosystem which denotes the cultural and political values and historical factors which shape the individual (Brofenbrenner, 1979; Lerner, 2002).

On the basis of these countless interactions within and across systems, all humans have the potential for a certain degree of systematic change (relative plasticity) across the lifespan. This plasticity suggests that the alignment of an individual (equipped with innate strengths) with a context replete with positive elements should result in beneficial outcomes (Lerner, 2005; Lerner, et al., 2003) and this forms the basis of the PYD philosophy.

**Developmental Assets and The 5 C’s of Positive Youth Development**

The Search Institute®’s Developmental Assets® Framework is the result of research investigating the nature of these positive individual and contextual elements which are presumed to be the key to a healthy transition to adulthood (Benson, 1997; Benson, Leffert, Scales, & Blyth, 1998; Leffert, et al., 1998; Lerner, 2002). The framework is comprised of 40 developmental assets divided into 20 internal and 20 external assets. The internal assets are further grouped into 4 categories: commitment to learning (e.g. academic motivation, completing homework, reading); positive values (e.g. helping others, being honest and demonstrating integrity); social competencies (e.g. resisting peer pressure,
planning and decision-making, cultural and interpersonal competence); and positive identity (e.g. personal control, self-esteem, positive expectations for the future; Benson, 1997; Benson et al., 1998; Leffert et al., 1998). Likewise, the 20 external assets fall into 4 groupings: support (e.g. from family, schools, and neighbours); empowerment (e.g. youth feel valued, safe and have useful roles); boundaries and expectations (e.g. clear rules and high expectations exist at home, school and in the community, positive adult role models are available, peers are a positive influence); and constructive use of time (e.g. participation in youth programmes, religious and creative activities, and time is spent with family; Benson, 1997; Leffert et al., 1998).

Supporting the value of these assets, empirical evidence indicates that the greater the number of assets a young person exhibits or experiences, the greater the prospect that he or she will display indicators of thriving (Scales, Benson, Leffert, & Blyth, 2000). Moreover, a study surveying nearly 100,000 American youth in grades 6 through 12 demonstrated that participation in risk behaviours was negatively associated with the accumulation of developmental assets (Leffert et al., 1998).

In a similar vein, a central assumption of the PYD philosophy is that an adolescent can be described as happy, healthy and functioning positively to the degree that they express 5 C’s: Competence (cognitive, behavioural and social abilities), Connection (healthy bonds with individuals and social institutions), Character (integrity and morality), Confidence (positive sense of self-worth and efficacy), and Caring (or Compassion- empathy and social justice values; Lerner, 2002, 2005; Lerner et al., 2003). The benefits of cultivating the 5 C’s extend beyond that of the individual because the combination of these attributes engenders the propensity to make positive contributions to family, the community and greater society in addition to positive self-contributions. Accordingly, this observation led to the designation of a sixth C: Contribution (Lerner, 2002, 2005; Lerner, et al., 2003). These concepts have been corroborated by results from a study by King et al.

---

4 In the Scales et al. (2000) study thriving indicators were specified as school success, overcoming adversity, maintaining physical health, and delaying gratification.
(2005) assessing whether the terms used by practitioners, parents, and youth to describe a “thriving” adolescent matched with the 6 C’s described in the research literature. Though the specific words used were not consistent across individuals in the study, researchers could easily categorise them within the 6 C groupings.

**Youth Development Programming**

Regrettably, the 21st century has been characterised by a deterioration in community participation and neighbourhood cohesion (Morrow & Styles, 1995; Putnam, 2001; Rhodes, 2005, Thompson & Kelly-Vance, 2001). People are no longer participating in work, religious, or volunteer organizations like they used to (Putnam, 2001) leaving people with fewer means of obtaining social support. The busy contemporary lifestyles of many parents also provoke beliefs that in addition to a lack of community ties, families are not as available to provide the resources young people need (Roth & Brooks-Gunn, 2003a, 2003b). Further, this lack of availability exists despite the quantity and severity of social problems not abating (Lerner, 2002) and adolescents of this generation experiencing a greater number of obstacles (e.g. a highly competitive employment market, and an increase in the availability and variety of illegal substances; Roth & Brooks-Gunn, 2003a, 2003b). It is not astonishing then that more than half of the American youth in the sample (n = 99,462) surveyed by researchers from the Search Institute reported having under half of the 40 developmental assets (Benson et al., 1998; Leffert et al., 1998) reducing the likelihood that they would reach their potential as young adults. Hence the main role of youth development programmes is to compensate for these deficits by providing an asset-rich environment which will promote the 6 C’s (Roth & Brooks-Gunn, 2003a, 2003b; Roth, Brooks-Gunn, Murray, & Foster, 1998).

**YOUTH DEVELOPMENT IN AOTEAROA NEW ZEALAND**

The relevance of a strengths-based approach to youth development has not escaped government policymakers in New Zealand. Growing concerns about young New Zealanders entering adulthood unprepared for the demands of a competitive job market and a greater prevalence of mental health problems in the
12-24 year age range, led the Ministry of Youth Affairs (MYA)\(^5\) to re-evaluate the current strategies available to support adolescents on their journey to adulthood (Ministry of Youth Affairs, 2002). Embracing the PYD approach, the Youth Development Strategy Aotearoa (YDSA) was launched in early 2002 to serve as a common platform for all sectors of government to use to inform youth-oriented policies and programmes. An MYA project team developed the strategy after an examination of the relevant research literature and consultation with youth-serving agencies and adults. The input of 1,450 New Zealand youth also played a role in shaping the strategy (Ministry of Youth Affairs, 2002).

The aims of the YDSA are to ensure that all government policies and practices comply with a PYD approach and to provide youth with access to a wide variety of PYD opportunities including connections to fundamental social networks. Six principles build the foundation for the YDSA: 1) Youth development is shaped by the big picture; 2) youth development is about people being connected; 3) youth development requires a consistent strengths-based approach; 4) youth development occurs through quality relationships; 5) youth development is triggered when youth fully participate; and 6) youth development requires reliable information. Following these principles should, in consequence, breed a nation of youth who have healthy bonds with others and feel a part of society; who are aware of the many options they have for a positive future; who feel that their contributions are valued; and who feel proud of who they are (McLaren, 2002; Ministry of Youth Affairs, 2002).

The “Big Picture” for Youth in New Zealand

As noted, the influence of multiple systems-level factors on adolescent development should not be understated (Lerner, 2002, 2005; Ministry of Youth Affairs, 2002); thus it is useful to delineate some of the macrosystemic influences affecting young people in contemporary New Zealand society, beginning with the historical-cultural context.

\(^5\) The Ministry of Youth Affairs and the youth-oriented functions of the Ministry of Social Development were amalgamated to form the Ministry of Youth Development in 2003 (http://www.myd.govt.nz/about-myd/index.html).
Historical Context

The establishment of New Zealand as a sovereign nation coincided with the signing of *Te Tiriti o Waitangi* /The Treaty of Waitangi on February 6th, 1840. This agreement between Māori and the British Crown aimed to give rise to a bicultural nation shared by two peoples, the tangata whenua (the people of the land) and the tangata tiriti (“the people whose presence was authorised by the Treaty of Waitangi”; King, 2007, p. 145). In signing the treaty, over 500 Māori chiefs agreed to cede sovereignty of the land to the Crown in exchange for protection and equal rights of British citizenship. Māori were also afforded *tino rangatiratanga* -- autonomous control of their resources and affairs (King, 2007; Orange, 2004). Māori perceived this to be a beneficial and equal partnership; the Crown would take responsibility for their compatriots and Māori would manage their own peoples and land alongside of the British government (Orange, 2004). What followed, however, was the cultural hegemony typical of this colonial era (McCarthy, 1997).

Traditional Māori society was based on kinship ties (Evans, Jory & Dawson, 2005; Pihama & Penehira, 2005). “The collective mode of operation as found in the whānau or extended family concept is fundamental to Māori culture.” (McCarthy, 1997, p. 27) and the whānau is still considered to be an essential contributor to health and positive identity development (M. H. Durie, 1985; McCarthy, 1997; Pihama & Penehira, 2005). The confiscation of Māori land by European colonialists and the resulting urban migration of Māori led to the alienation of whānau from each other, their *hapū* (subtribe), *iwi* (tribe), and their land (Puketapu-Andrews, 1997). The disruption of these connections, the suppression of Māori spiritual beliefs and language, and the general derogation of the indigenous culture by the non-indigenous settlers resulted in the deterioration of a positive personal and collective identity for Māori (A. Durie, 1997). Additionally, the implementation of a substandard educational curriculum for Māori severely limited their employment and earning potential (Jahnke, 1997; Orange, 2004). This kind of institutionalised racism (see Jones, 2000, 2001) set the stage for the disparities between Pākehā and Māori in educational, economic and health spheres that are still evident today (see...
The establishment of the Waitangi Tribunal in 1975 marked a positive turn for indigenous rights in New Zealand. The tribunal provided a vehicle through which Māori could submit claims of injustices by the Crown for inquiry (Orange, 2004). Additionally, there has been general public acknowledgement that the Treaty was breached and support for the restoration of Māori tino rangatiratanga and the bicultural society envisioned by the Treaty is growing. Government policies and practices are now obligated to reflect the principles of the Treaty (Orange, 2004; Te Röpū Rangahau Hauora a Eru Pōmare, 2002). Accordingly, the YDSA endorses an approach to the development of young Māori that is underpinned by Māori values and supports the restoration of ties to their whānau, hapū, and iwi (Ministry of Youth Affairs, 2002).

A Description of the New Zealand Youth Context Today

The general population of New Zealand is approximately 4.4 million (Statistics NZ, 2006), and young people aged 10–19 years form approximately 15% of that figure. The breakdown of the main ethnic groups within this 15% is estimated to be: 62% European, 21% Māori, 10% Pacific, and 10% Asian (Statistics NZ, 2006) and the younger demographic is becoming increasingly diverse (Dunphy et al., 2008; Ministry of Social Development, 2010).

The first national survey assessing the health and well-being of secondary school students in New Zealand was conducted in 2001. An updated version of the survey was administered to another cohort of over 9,000 randomly selected youth in 2007. The Youth’07 survey reflects the most holistic and up-to-date picture of secondary school student health and well-being in the country and provides the first point of comparison for investigating health and well-being trends (Adolescent Health Research Group, 2008). Overall, the comparative results show improvement.

In contrast to the 2001 cohort, participants in 2007 reported better relationships with peers and family; greater school engagement, connectedness,
and safety; greater participation in youth groups; improvements in eating behaviours and physical activity; reductions in most risk behaviours and suicide rates; and increases in emotional well-being and life satisfaction (Adolescent Health Research Group, 2008). Interestingly, these marked improvements across educational, relationship, and health domains occurred during the years just following the dissemination of the YDSA.

On the other hand, alcohol consumption is still high (61%) and a considerable proportion take part in regular binge drinking (34%), risky sexual practices have not improved, sedentary behaviour has increased, as has the witnessing of violence in the home (6% in 2001 to 10% in 2007). Concerns about the proportion of families experiencing economic hardship still remain. Approximately one tenth of the sample reported substantial levels of economic adversity.

In the general New Zealand population, individuals from highly deprived areas\(^6\) report poorer health, lower educational qualifications, and more instances of criminal victimisation (Ministry of Social Development, 2010). Young people from neighbourhoods of high deprivation in the Youth’07 sample also reported poorer health and well-being across several indicators in comparison to those living in low deprivation contexts (Adolescent Health Research Group, 2008). Furthermore, Māori are generally over-represented in areas of high deprivation (Clark, et al., 2008; Ministry of Social Development, 2010). Pacific peoples have a similar pattern of over-representation in low socioeconomic communities (Helu, Robinson, Grant, Herd, & Denny, 2009; Ministry of Social Development, 2005) and also experience worse health and well-being outcomes than the general New Zealand population (Ministry of Social Development, 2010; Ministry of Youth Affairs, 2002).

Like Pākehā, most Māori and Pacific students report good health, caring relationships, and positive school environments; however, Māori and Pacific youth

\(^6\) Deprivation scores are measured by the NZDep2006 index which uses income, home ownership, employment status, educational qualifications, living space, and communication and transportation constraint data to categorise geographical areas termed meshblocks according to a scale from 1 (low) to 10 (high) (Salmond, Crampton, & Atkinson, 2007).
report more mental health issues and substance use; less consistent use of contraception; reduced access to health services; and greater exposure to violence in the home than Pākehā youth (see Clark et al., 2008 for the Youth’07 report on Māori youth and Helu et al., 2009 for the data pertaining to Pacific peoples). Māori adolescents are also less likely to report high expectations from their teachers (Clark et al., 2008) and Pacific students are less likely to report that their teachers treat them fairly (Helu et al., 2008) and these two groups are more likely to leave secondary school with lower educational qualifications than their Pākehā counterparts (Ministry of Social Development, 2010).

In recent years, however, a larger number of students left school with higher qualifications and this increase is more pronounced for Māori and Pasifika groups (Ministry of Social Development, 2010). Young Māori and Pacific youth in the Youth’07 sample indicated they had lower substance use and better mental health than those in the 2001 sample. Additionally, both Māori and Pacific adolescents reported a greater sense of school belonging than Pākehā adolescents (Clark et al., 2008; Helu et al. 2009).

Asian youth in New Zealand fare better than the overall New Zealand youth population when it comes to obtaining higher educational qualifications. They also have lower participation in risk behaviours in general (Ministry of Social Development, 2010). In 2004, however, Chinese and Indian students (the largest groups within the Asian category) reported more symptoms of depression that young Pākehā and Asian adolescents in general reported other unique concerns. For example, the majority felt misunderstood by their families; many did not feel safe at school; and although less instances of bullying were reported, the experiences were more traumatic than those of Pākehā students (Rasanathan et al., 2006). Asians, as an overall group, are also more likely to feel personally discriminated against in comparison to the other ethnic groups in New Zealand (Ministry of Social Development, 2010).

---

7 At the time of writing, no report for young Asians was available from the Youth’07 survey.
In summary, although the majority of young people in New Zealand report good general health and positive family, peer, and school environments, a number of social and health concerns are evident across all groups of New Zealand adolescents. On the whole, this contextual analysis provides sufficient justification for the maintenance and creation of asset-rich settings that nurture positive adolescent growth. Additionally, the ethnic differences described above suggest that young people belonging to Māori, Pacific, and Asian ethnic groups have unique needs that should be considered. Jones (2000, 2001) argues that ethnic disparities in well-being outcomes are generally linked to reduced exposure and access to the social determinants of health as a result of continued exposure to institutionalised, personally mediated, and internalised racism and the findings above imply that racial factors may in fact underlie some of the discrepancies found between ethnic groups in the New Zealand youth population. Evidently, a more in-depth exploration of how asset-rich settings fit with the unique values and experiences of different ethnic groups is needed.

THE VALUE OF YOUTH PROGRAMMES

Evidence exists to justify the importance of youth programmes in promoting positive development. Scales et al., (2000) demonstrated that time spent in youth programmes was one of most effective predictors of a composite index of thriving indicators (based on the Search Institute®’s 40 Developmental Assets®) across all of the racial-ethnic groups included in the study. Additionally, when assessing the predictors of each individual thriving indicator, involvement in youth programmes had the most pervasive effect: influencing five of the seven indicators measured (Scales et al, 2000).

Not surprisingly, the 5 (or 6) C’s of the PYD approach are often used in stated outcomes of youth programmes (Lerner, Fisher, & Weinberg, 2000; Nicholson, Collins & Holmer, 2004; Roth & Brooks-Gunn, 2003a, 2003b) and a review of 48 youth programme evaluations demonstrated that of the 23 programmes measuring competency outcomes, 21 were successful in effecting improvements. All 19 programmes targeting connectedness outcomes exhibited
positive changes as did all programmes measuring character and caring goals (Roth & Brooks-Gunn, 2003b).

In another review of 25 American youth development programmes, Catalano et al. (2004) found evidence of changes in both positive and negative outcomes including improvements in one or more of: relationships, interpersonal and cognitive skills, self-efficacy, academic motivation and decreases in one or more of: substance use, misconduct at school, antisocial behaviour and risky sex practices. It should be noted that these 25 programmes were likely exemplars in the field as they were able to attract enough funding to implement a high quality evaluation. However, Durlak et al.’s (2007) more inclusive meta-analysis included 79 of 336 programmes that fell within the description of youth development and targeted changes in a young person’s ecological environment. This study produced impressive evidence that youth development programmes can also be effective in making a positive difference in relations at the systems level by changing family and school conditions.

THE FEATURES OF YOUTH DEVELOPMENT PROGRAMMES

Not all youth-serving programmes can be characterised as youth development programmes. Distinguished from programmes that primarily provide a safe space for unstructured recreation (Riggs & Greenberg, 2004) or to fill idle time (Roth & Brooks-Gunn, 2003a), all youth development programmes have an overall aim of enhancing programme participant strengths (Durlak et al., 2007; Gallagher, Stanley, Sherer, & Mosca, 2005; Roth et al., 1998) even if a simultaneous goal is to prevent problems (Roth & Brooks-Gunn, 2003a). Beyond this overarching purpose, a concrete operationalisation is difficult to pin down since programmes are very diverse (Gallagher et al., 2005; Riggs & Greenberg, 2004; Roth & Brooks-Gunn 2003a; Roth & Brooks-Gunn 2003b; Roth et al., 1998). Programme designs range from mentoring, education and training, and community service (Gallagher et al., 2005; Ministry of Youth Development, 2009) to adventure, and arts and culture programmes (Ministry of Youth Development, 2009); the duration and intensity of programmes differ, as do the characteristics of the participants (Riggs &
Greenberg, 2004) and the numbers served (Roth & Brooks-Gunn, 2003a).
Additionally, many programmes falling under the guise of prevention
programmes share similar characteristics to those identifying as youth
development programmes (Catalano et al., 2004; Durlak et al. 2007; Roth & Brooks-
Gunn, 2003b). This creates difficulty in deciphering which aspects of each
programme produce the positive effects (Roth & Brooks-Gunn, 2003a, 2003b).

Several researchers (Gallagher et al., 2005; Roth & Brooks-Gunn, 2003a,
2003b; Urban, 2008) have endeavoured to further define the distinguishing features
of a youth development programme by surveying the views of policy makers,
researchers, adolescents (Urban, 2008) and practitioners (Gallagher et al., 2005;
2003b) suggested that youth development programmes can be differentiated from
other youth programmes according to three main dimensions: goals, activities, and
atmosphere.

Goals can generally be classified as one or more of the 5 C’s (Lerner, 2002;
Roth & Brooks-Gunn, 2003a, 2003b) and the most pervasive goal is one of building
strengths or competencies (Catalano et al., 2004; Gallagher et al., 2005; Roth &
others is another common characteristic (Catalano et al., 2004; Gallagher, et al.,
is fostering a positive identity and self-determination (Catalano et al., 2004). Many
programmes have simultaneous problem prevention goals (Gallagher et al., 2005;
(2003b) also found that youth development programmes, over other types of
youth-serving programmes, more often target changes in a young person’s context.
Still, there is some debate as to whether it is more important to target individual
outcomes or to strive to change the young person’s environment (Roth & Brooks-
Gunn, 2003b).

In line with a goal of building competencies, skill-building activities are the
most prevalent (Roth & Brooks-Gunn, 2003a; 2003b; Roth et al., 1998). These may
include life skills (Urban, 2008), leadership skills (Roth & Brooks-Gunn, 2003a; 2003b; Urban, 2008), social skills (Catalano et al., 2004) and those of an academic or vocational nature (Roth & Brooks-Gunn, 2003b). New activities that expand the young person’s opportunities and/or worldview are also seen to be important (Catalano et al., 2004; Roth & Brooks-Gunn, 2003a). Likewise, challenging tasks that require active participation, (Larson, 2000; Roth & Brooks-Gunn, 2003a, Roth et al., 1998), authentic acts (Roth & Brooks-Gunn, 2003b) and opportunities for prosocial involvement (Catalano, et al., 2004), for example community service, help to further the programme goals.

The programme atmosphere is thought to hold the most significant weight in differentiating successful youth development programmes from other programmes for youth (Heinze, Jozefowicz, & Toro, 2010; Roth & Brooks-Gunn, 2003b). A youth-centred focus entails giving youth a voice in programme design and operations (Gallagher et al., 2005; Roth et al., 1998), although there is not necessarily consensus as to whether youth should be involved in all aspects of programme planning (Gallagher et al., 2005; Urban, 2008). This vital atmosphere of empowerment (Heinze et al., 2010; Roth & Brooks-Gunn, 2003a, 2003b; Urban, 2008) can also be reinforced by giving youth meaningful roles and responsibilities, and choices (Roth & Brooks-Gunn, 2003a; Roth et al., 1998). Youth agency should also be encouraged in the community (Nicholson, et al., 2004). In order to cultivate initiative and intrinsic motivation, adults need to scaffold new learning by keeping the young person on track while still supporting his or her autonomy (Larson, 2006).

Indeed, the quality of relationships with staff and other adults in the programme can also have a large impact on programme effectiveness (Grossman & Bulle, 2006; Heinze et al., 2010), hence programmes should be structured to be supportive of these relationships (Grossman & Bulle, 2006; Roth & Brooks-Gunn, 2003a, 2003b; Roth et al., 1998). Responsive and inclusive staff members pay attention to an individual’s unique characteristics but are at the same time sensitive to cultural backgrounds (Durlak et al., 2007; Lerner, 2002; Riggs & Greenberg; Roth & Brooks-Gunn, 2003a; Roth et al., 1998; Urban, 2008) and challenges in the young
person’s environment (Nicholson et al., 2004). Consequently, participants will more likely develop a strong sense of belonging (Heinze et al., 2010; Roth & Brooks-Gunn, 2003a; Roth et al., 1998; Urban, 2008). Positive behaviour and accomplishments should be recognised by adults (Catalano et al., 2004; Roth & Brooks-Gunn, 2003a, 2003b) but in accordance with high expectations (Roth & Brooks-Gunn, 2003a, 2003b) and prosocial norms (Catalano et al., 2004; Heinze et al., 2010).

Structure and clear rules within an agency has been acknowledged by researchers to be an important feature (Riggs & Greenberg, 2004; Roth et al., 1998). Homeless and high risk youth participating in one programme also reported that they were more satisfied when the programme was well structured and organised. This was especially pertinent to the older youth in the programme (Heinze et al., 2010); however, other researchers feel that this is secondary to other characteristics (Urban, 2008). Safety (Heinze et al., 2010; Roth et al., 1998; Lerner, 2002), both a physical and psychological, is a necessity (Roth & Brooks-Gunn, 2003a). Family (Heinze et al., 2010) and community involvement (Urban, 2008; Lerner, 2002) are also believed to improve results but these are mentioned less often than the above characteristics.

Additionally, programme duration can make an important difference, with programmes lasting longer than nine months (Catalano et al., 2004), or at least the duration of the school year (Roth & Brooks-Gunn, 2003b; Roth et al., 1998), producing better outcomes. In their reviews, Catalano et al. (2004) and Roth et al. (1998) noted that programmes incorporating more of the above characteristics demonstrated the most success.

**THE STATE OF YOUTH PROGRAMME EVALUATION STUDIES**

According to Roth et al. (1998), the state of youth programme evaluation research is in its infancy with studies only arising in the late 1990s (Roth et al., 1998), leaving a limited pool of evidence to base effectiveness conclusions on. The studies that have been conducted often lack in methodological rigor; few
evaluations employ experimental or quasi-experimental designs thus offering no adequate comparison group (Catalano et al., 2004; Riggs & Greenberg, 2004; Roth et al., 1998; Roth & Brooks-Gunn, 2003b). Follow-up studies are also scarce but necessary to demonstrate sustained effects (Catalano et al, 2004).

Furthermore, few programmes incorporate most of the important features. Roth & Brooks-Gunn’s (2003b) review of 48 youth programmes found that only two of these programmes met the criteria of targeting the broad goals manifested by the 5 C’s, providing a supportive and empowering atmosphere conducive to skill-building and authentic activities that broaden youth horizons, and was of a sufficient duration. Positive results do exist but occasions of mixed, null and even harmful effects have also appeared in the literature (Riggs & Greenberg, 2004).

There is also concern that the exclusive focus in outcome evaluations on whether or not the programme works is not suitable for answering valuable policy and programme development questions (Riggs & Greenberg, 2004). A much needed focus on programme processes, contextual factors and individual differences in evaluation research is needed (Larson, 2000; Riggs & Greenberg, 2004; Roth & Brooks-Gunn, 2003a, 2003b; Roth et al., 1998). Greater use of mixed-method designs which draw on the strengths of both quantitative and qualitative approaches would be useful for developing a more nuanced understanding of programme effects, as would mediation and moderation techniques (Riggs & Greenberg, 2004) and longitudinal studies (Larson, 2000). This would help us address pivotal questions such as: Why does the programme work (Roth & Brooks-Gunn, 2003a, 2003b); which elements are the main contributors to programme success (Roth et al., 1998); under what conditions do they achieve the best results (Riggs & Greenberg, 2004) and for which types of individuals do they work for (Larson, 2000, Riggs & Greenberg, 2004; Roth et al., 1998)?

In New Zealand, the situation is even more concerning because there is a paucity of good quality evidence on youth programming. In the interests of developing evidenced-based policy and practice, the Ministry of Youth Development reviewed the best practice and effectiveness evidence for structured
youth development programmes but much of the information was based on research from overseas as the evidence base in New Zealand is so limited (Ministry of Youth Development, 2009).

Taking a closer look at the programme designs that are of particular relevance to Project K (adventure programmes, community service-learning, and mentoring) can contribute to knowledge about programme effectiveness as well as the above-mentioned pivotal process questions. In view of this, the next sections focus on youth development through adventure programming, service-learning and, lastly, mentoring.
OUTDOOR ADVENTURE PROGRAMMES

Great things happen when Men & Mountains meet. This is not Done by Jostling in the Street

Challenge is what makes men. It will be the end when men stop looking for new challenges
~Sir Edmund Hillary.

The feeling of adventure is, for some, the means to the optimal subjective experience (Csikszentmihalyi & Csikszentmihalyi, 1990). In terms of personal development, even the ancient Greeks recognised the potential that high-risk activities could have in building the physical and emotional faculties of young Spartans (Ewert, 1989). The opening quotes epitomise the essence of outdoor adventure programmes. Blake’s rhyming couplet encapsulates the idea that encounters with Mother Nature can provide a true test of human potential and Hillary’s statement captures the “development-by-challenge” philosophy (Neill & Dias, 2001, p. 35) that lies at the heart of adventure-based initiatives.

THE ORIGIN AND GROWTH OF OUTDOOR ADVENTURE PROGRAMMING

Perhaps the first version of adventure education as we know it today was Outward Bound (OB), established in Wales in 1941 by Kurt Hahn in conjunction with Lawrence Holt (Priest and Gass, 2005). Typical OB programmes included challenging teamwork activities in the outdoors to develop technical skills; an expedition; one to three day solos, a service project, an endurance challenge, and a celebration, and these are still the standard features of many modern programs (Hirsch, 1999).

The impetus behind the inaugural OB programme has been attributed to the need to better prepare young men in Britain’s Merchant Navy for the perils of the Second World War (Greene & Thompson, 1990; Hopkins & Putnam, 1997) but the programme served young men from a variety of backgrounds (Miner, 1990; Priest & Gass, 2005). Lawrence Holt’s frequently noted phrase that the training was “less a training for the sea than a training through the sea” (Hopkins & Putnam, 1997, p.
suggests that it was more likely an opportunity to use the natural elements to spread the values of leadership, service, physical preparedness and active citizenship that provided the foundation for the innovative curriculums of Hahn’s earlier Salem (Germany) and Gordounstoun (Scotland) schools (Priest & Gass, 2005; Richards, 1990). These were values Hahn believed needed to be instilled to combat the societal ills brought on by an increasingly industrialised world. He conceptualised these ills as six social declines: the decline of fitness, of initiative and enterprise, of memory and imagination, of skill and craftsmanship, of self-discipline, and of compassion (Hopkins & Putnam, 1997; Priest & Gass, 2005; Richards, 1990).

The 1950s to the 1980s was a period of rapid growth in the adventure education field (Hopkins & Putnam, 1997). Branches of OB opened in several different countries and spawned spin-off programmes (Miner, 1990). The OB model was applied to specialised populations of delinquent adolescents as early as the 1960s. Later this evolved into a model that combined the traditional adventure curriculum with therapeutic strategies (Russell, 2001). Other modifications to the original model allowed it to be integrated into the U.S. high school curriculum (Priest & Gass, 2005; Prouty, 1999) and this then generated adaptations for students with special needs, clinical populations, and corporate settings (Prouty, 1999). In 2009, OB programmes existed in 34 countries serving more than 250,000 participants a year (Outward Bound International, 2009) and many more are served through other programmes following the same philosophy.

OUTDOOR EDUCATION, ADVENTURE PROGRAMMES, WILDERNESS THERAPY – ONE AND THE SAME?

The various derivatives of the traditional OB model mean one is likely to confront a myriad of terms describing adventure-based endeavours (Hattie, Marsh, Neill, & Richards, 1997; Russell, 2001). For example, adventure programmes (Priest & Gass, 2005), adventure education (Prouty, Panicucci, & Collinson, 2007), wilderness programmes (Orren & Werner, 2007; Romi & Kohan, 2004), wilderness therapy (Russell, 2001), adventure-based programming (Shirilla, 2009; Whittington & Mack, 2010), adventure-based counseling (Prouty, 1999; Schoel, Prouty, &
Radcliffe, 1988), bush adventure therapy (Pryor, Carpenter & Townsend, 2005),
and outdoor adventure pursuits (Ewert, 1989) are some of the labels that have been
applied to these initiatives. Yet the two major divisions that should be noted within
the broader field of outdoor education, are the environmental education branch which
draws explicit attention to relationships with and within nature (Priest, 1999c;
Priest & Gass, 2005) and the adventure education (or adventure programming) branch
in which the betterment of individuals and human society are the end goals. The
‘great outdoors’ (or simulated environments) simply provide an optimal setting for
these processes to occur (Priest, 1999c; Priest & Gass, 2005). What defines these
programmes as adventure-based is that they all purposefully utilise elements of
uncertainty and risk for positive ends (Ewert, 1989; Gregg, 2007; Hirsch, 1999;
Horwood, 1999; Priest, 1999c; Prouty, 2007; Priest & Gass, 2005). They also
typically serve small groups of participants and, as the higher order outdoor
education classification would imply, activities mainly occur in the outdoors
(Pryor et al., 2005).

RATIONALE FOR ADVENTURE PROGRAMMING

Hopkins and Putnam (1997) argue that the need for Hahnian methods and
values has not abated in the current social climate. The deterioration of physical
fitness and the prevalence of obesity is acknowledged to be a problem of pandemic
proportions (Roth, Qiang, Marban, Redelt, & Lowell, 2004). Hopkins and Putnam
(1997) noted the relevance of this particular decline in today’s society and also
expressed dismay over the personal and societal consequences of rising
unemployment rates. They suggested that the information overload characteristic
in current society hinders and Hahn’s ideal concept of citizenship is not realised
because the current focus is on obtaining individual rights rather than making
contributions to society.

While Hopkins and Putnam (1997) did not concur with Hahn’s claim of a
decline in initiative and enterprise either in his generation or in contemporary
society, research by Larson and his colleagues suggest otherwise (Larson, 2000;
Larson & Richards, 1991). For example, in a study with 392 Chicago-based
adolescents, Larson and Richards (1991) found a substantial portion of the participants’ time in and out of school was characterised by boredom. Larson (2000) explained that the development of initiative necessitates conditions that are intrinsically motivating and engage attention – quite the antithesis of boredom. He also asserted that the existence of initiative can create flow-on effects of increased leadership, altruism and civic engagement. Interestingly, these are precisely the characteristics Hahn was attempting to cultivate through his programmes. An assessment of the change processes thought to be operating within these programmes to breed such qualities follows.

**PROCESSES OF CHANGE**

**Experiential Learning Theory**

John Dewey (1963) condemned pedagogical practices that were disconnected from future experiences. The progressive education movement he promoted centred on a philosophy of teaching through educative experiences that engaged the learner and directly influenced subsequent experiences. These beliefs were the seeds of the experiential learning theories that undergird adventure programming (Ewert & Garvey, 2007; Hirsch, 1999; Hopkins & Putnam, 1997; Luckner & Nadler, 1997; Panicucci, 2007; Priest & Gass, 2005).

David Kolb, another name frequently linked to experiential learning (Panicucci, 2007; Wurdinger & Priest, 1999), drew together key aspects from the theoretical perspectives of John Dewey, Kurt Lewin, and Jean Piaget to describe the nature of experiential learning theory (Kolb, 1984). A theory that, by giving the learner an active and central role in his or her learning, highlights the importance of giving personal meaning to abstract concepts (Luckner & Nadler, 1997), thus grounding lessons in reality (Hirsch, 1999; Wojcikiewicz & Mural, 2010) and making clear the connections between current actions and future consequences (Wojcikiewicz & Mural, 2010). In contrast to the cognitive view of learning which emphasises the manipulation and recollection of abstract symbols and behavioural theories which construe learning in terms of stimulus-response with no mediating
cognitive or experiential roles, Kolb (1984) argued that experiential learning is holistic. It engages the physical, intellectual, emotional (Ewert, 1989; Hopkins & Putnam, 1997; Luckner & Nadler, 1997), and social facilities (Luckner & Nadler, 1997) and the use of all senses (Kraft, 1999; Priest, 1999c). Consequently, this method of learning is more engaging for the learner as it instils a sense of ownership (Luckner & Nadler, 1997). Knowledge is more thoroughly internalised and believed to be more effective in developing the multiple dimensions of intelligence (Kraft, 1999).

Kolb further discredited cognitive and behavioural learning theories for being de-contextualised. He contrasted these with learning through experiential means during which the learner is obviously intimately connected to the context. Kolb (1984) suggested that learning arises from transactions with the context and explicates that a transaction (as opposed to an interaction) implies that both the person and the environment are reciprocally changed as a result. Kolb also (1984) described experiential learning as a continuous process, generally conceived as four cyclically iterated phases (Hopkins & Putnam, 1997; Luckner & Nadler, 1997; Panicucci, 2007). The cycle begins with the learner participating in a concrete experience (1); after reflecting on the experience (2); he or she makes abstract generalisations (3) that can be extrapolated to similar situations; the learner actively experiments (4) with the hypotheses just made by applying them in a new concrete experience (1); and so, the cycle begins again (Hopkins & Putnam, 1997; Kolb, 1984; Luckner & Nadler, 1997; Panicucci, 2007; Priest & Gass, 2005).

**Experiential Learning and Outdoor Adventure Programmes**

Walsh and Golins (1976) put forth a theoretical model detailing the specific change process thought to be operating in the adventure programme context: The Outward Bound Process. Their model proposed adventure programmes can be effective in reorganising the meaning of a participant’s life and its future direction because the nature of the physical and social environment and the characteristics of the required problem-solving activities motivate the participant to master the tasks at hand. Despite a dearth of empirical evidence to support the process, this
description is the most widely accepted understanding of what transpires during structured outdoor adventure experiences (McKenzie, 2003; Sibthorp, 2003a). These aspects and the importance of the adventure programming instructors in facilitating participant growth are elaborated on below, with support from other studies.

**Physical Environment**

The critical aspect of the physical environment is that it is unfamiliar (Luckner & Nadler, 1997; McKenzie, 2000; Priest & Gass, 2005; Walsh & Golins, 1976). The disparity between the novel surroundings and the conditions participants are typically accustomed to creates a sense of internal tension or disequilibrium which engages the participant’s attention allowing them to notice things they may neglect to perceive in their normal environment (Luckner & Nadler, 1997; McKenzie, 2000; Walsh & Golins, 1976). Duly, adventure programme participants have reported that the novelty of the environment does influence individual transformation (Garst, Scheider & Baker, 2001; Lynch, 2000; McKenzie, 2003; Smith, Steel, & Gidlow, 2010). Wilderness settings are thought to be particularly effective in engendering a sense of risk and unpredictability that foster the necessary state of dissonance (Miles, 1999). An outdoor environment provides consequences that cannot be avoided (McKenzie, 2000; Walsh & Golins, 1976); thus, generating a motivating force to take action.

**Social Environment**

For some, the novelty of the context is found in the way the physical constraints of the situation influence social interactions (Beames, 2004; Eggleston, 2000; Sibthorp, 2003b). Isolation in the wilderness with a small group of people and no chance of escape can create an atypically intense social experience. Hopkins and Putnam (1997) likened this to an “experimental social laboratory” (p. 13) where participants can assume different social roles and explore new behaviours while being able to clearly see the consequences of their actions. Conditions such as these may fuel mild interpersonal conflict but a bit of conflict under emotionally safe conditions is an important part of the process (Quay, Dickinson, & Nettleton, 2000; Ewert & Heywood, 1991; McKenzie, 2000; Quay et al., 2000; Walsh & Golins, 1976).
It is through conflict resolution that people gain new perspectives of others, which is key to a caring community-oriented climate; therefore, it may even be purposely instigated by the facilitators (Ewert & Heywood, 1991; Quay et al., 2000).

The ideal group size is between seven and fifteen people. Diverse perspectives are bound to exist within a group of this size and initially this may increase the likelihood of conflict yet it should be small enough that factions within the group do not develop and conflicts can be worked out (McKenzie, 2000; Walsh & Golins, 1976). Furthermore, the explicit goals of adventure programmes are such that interdependence is required to attain them (Hirsch, 1999) thus the resolution of the conflict is usually inevitable (McKenzie, 2000; Quay et al., 2000; Walsh & Golins, 1976). Participants involved in an overseas adventure education programme with a group of individuals unknown to them previously recognised that being forced to address conflict within the group was important and, in time, the diverse perspectives of the varied members broadened their own worldviews (Beames, 2004).

Once adjusted to the unique social conditions, strong bonds between group members usually develop. Structured opportunities for cooperative tasks that require reliance on others (Garst et al., 2001; Martin & Leberman, 2005; McKenzie, 2003) and open communication (Luckner & Nadler, 1997) expedite the bond-building process. The positive relationships that result can be considered positive outcomes in and of themselves but the support a group provides during the experience can also enhance the probability of individual achievement in the programme (Hirsch, 1999; Hopkins & Putnam, 1997). Neill and Dias (2001), for example, found that, for a group of young adults involved in Australian OB courses, the greater the level of support offered from their group’s least supportive member, the larger the increase in the group’s mean resilience scores from pre to post-programme.

A supportive group context (Eggleston, 2000; Martin & Leberman, 2005; Martin & Legg, 2002), teamwork activities (Garst et al., 2001; Martin & Leberman, 2005; McKenzie, 2003), and the general social learning opportunities (Sibthorp,
2003b) are factors participants have also confirmed as influencing their growth. Conversely, one study by Sibthorp and Arthur-Banning (2004) showed that group functioning was not a significant predictor of overall life effectiveness. Group dynamics may not always be directly linked to all of the targeted outcomes but numerous qualitative studies (e.g. Beams, 2004; Eggleston, 2000; Martin & Leberman, 2005; Martin & Legg, 2002; McKenzie, 2003; Sibthorp, 2003b) suggest that a cooperative group is integral to programme-induced social growth.

**Activities**

Shellman and Ewert (2010) concluded that it was the sense of achievement participants felt during a U.S.-based OB programme that led to an increase in their general sense of psychological empowerment. People are not likely to feel this sense of achievement if they remain in their comfort zone. Luckner and Nadler (1997) explain that growth and learning occur as a result of pushing past the “comfort zone” to a “groan zone” (p. 20) where one needs to apply their competencies to resolve the discomfort and once the tension has been alleviated, a feeling of mastery and confidence follows. Therefore, another source of dissonance that motivates action is the challenging nature of outdoor adventure activities (Ewert, 1989; Luckner & Nadler, 1997; McKenzie, 2000; Walsh & Golins, 1976) – and participant reports also substantiate this essential element (Martin & Leberman, 2005; Martin & Legg, 2002; McKenzie, 2003).

The challenge should provoke an uncertainty of success but in the end achievement needs to be attainable. The most effective way to ensure this is to organise activities in a sequential manner such that each task is incrementally more difficult than the previous (Hirsch, 1999; McKenzie, 2000; Walsh & Golins, 1976). A recognisable beginning and end to an activity is important for gauging achievement so concrete tasks are more appropriate (McKenzie, 2000; Walsh & Golins, 1976) as are authentic ones (Hopkins & Putnam, 1997; Sibthorp, 2003b). Authentic activities generate clear-cut consequences that are experienced physically, emotionally, and intellectually (Walsh & Golins, 1976).
Reflection and feedback obtained after an activity allows participants to internalise the meaning of the experience (Hopkins & Putnam, 1997) and so activities that focus on processing the experience (e.g. group debriefs, time to reflect alone, journal writing) are essential to the experiential learning cycle. Setting time aside to process the experience as a group provides the opportunity to obtain valuable feedback, in addition to the feedback obtained when observing others modelling appropriate behaviours or while an individual is participating. For educational interventions as a whole, feedback has actually been identified as the most important moderator of achievement (Hattie, 1992). Accordingly, individual growth has been linked to having group discussions and taking part in a solo (Gassner & Russell, 2008; McKenzie, 2003).

Instructors

Granted that instances exist where instructor support has not emerged as a significant predictor of adventure education outcomes (Gassner & Russell, 2008; Sibthrop & Arthur-Banning, 2004), when asked, participants have on several occasions acknowledged it to be an important factor in their personal success (Eggleston, 2000; Martin & Leberman, 2005).

Contrary to the belief that outdoor experiential education simply entails transporting individuals to unconventional locales and leaving them to their own devices to extract meaningful knowledge from their experiences, the instructor plays an important and purposeful role in facilitating learning. Itin (1999) distinguishes experiential learning from experiential education by suggesting that the former is an internal individual process while the latter relies on transactions between an instructor and a student over the same concrete experience. An instructor is responsible for identifying and constructing opportunities for experiential learning to occur, providing pertinent information, setting the boundaries of the experience, and guiding learners through the reflection and generalisation stages (Itin, 1999; Luckner & Nadler, 1997) – including drawing comparisons with possible future situations to facilitate the transfer of knowledge (Gass, 1999; Luckner & Nadler, 1997). Competency in teaching and facilitation also means having to understand the benefits and pitfalls of different pedagogical
strategies and the appropriate time and place to enact these (Luckner, 1994) which coincides with a flexible leadership style (Priest, 1999b; Priest & Gass, 2005).

The nature of the adventure activities and the outdoor setting require a much wider complement of skills, however. Abseiling, kayaking, rock climbing and other typical outdoor pursuit activities all necessitate proficiency in specific technical skills; ensuring participant safety requires first aid skills and knowledge of environmental conditions and emergency procedures (Priest, 1999b; Priest & Gass, 2005; Prouty et al., 2007); and relating and communicating well with others is essential in order to resolve intergroup conflicts and to provide empathetic support. In addition, the adventure programme leader must be organised, have a keen sense of judgement, and be adept at problem-solving (Priest, 1999b; Priest & Gass, 2005; Prouty et al., 2007).

The most crucial problem to be solved in situations such as these is, arguably, “the problem of the match” (Hopkins & Putnam, 1997, p.102). To maximise the feelings of adventure, accomplishment and, consequently, empowerment, instructors must be able to construct experiences that provide the optimal level of stimulation (Hopkins & Putnam, 1997). This involves striking a balance between risk and excitement and an individual’s actual level of competency (Priest, 1999a; Priest & Gass, 2005). To ensure physical safety, an instructor needs to play up the perceived risk of the situation while maintaining a low level of real risk (Ewert, 1989; Hopkins & Putnam, 1997; Priest, 1999a; Priest & Gass, 2005).

Perceived risk is a subjective perception and therefore will vary from individual to individual (Gregg, 2007) as does competency level. An effective programme leader needs to be attuned to individual differences if he or she is to maximise the individual’s learning (Hopkins & Putnam, 1997; Priest, 1999; Priest & Gass, 2005). In the most benign form, inaccuracy in this critical skill can lead to mere experimentation and exploration that does not increase the young person’s skill or motivation (Priest 1999a). At worst, the existence of real risk in adventure-based endeavours (Ewert, 1989; Gregg, 2007; Horwood, 1999; Hirsch, 1999; Priest,
1999a; Prouty, 2007) means that inaccurate judgements can lead to some serious deleterious effects. The recent school canyoning tragedy, where six students and one teacher lost their lives while on an outdoor pursuits expedition in New Zealand, (Binning, 2008) makes this fact glaringly obvious.

Dangerous and fatal incidents provoke quick responses to increase the physical safety standards in programme procedures (Davidson, 2004; Lynch, 2006) and improve instructor training (Lynch, 2006) but Davis-Berman and Berman (2002) contend that more attention is needed to mitigate the potential emotional or psychological consequences that can arise from involvement in adventure courses. Pushing people too far beyond their comfort level can provoke extreme anxiety (Davis-Berman & Berman, 2002; Ewert, 1989; Prouty, 2007) and lead to persistent traumatic effects (Davis-Berman & Berman, 2002). Fabrizio and Neill (2005) illustrate how being placed in a foreign environment, whether it is in a new country or novel wilderness surroundings, can induce symptoms of culture shock. When supported in the right way the experience of culture shock may enhance cultural adaptation skills for the future but if the right precautions are not taken it may have enduring negative emotional effects (Fabrizio & Neill, 2005). In accordance with this, instructors need to be attuned to an individual’s emotional as well as physical limits. The potential for disastrous consequences makes an assessment of each adventure programme’s merit or worth all the more important.

**THE SUCCESS OF ADVENTURE PROGRAMMES**

According to Larson (2000), adventure programmes have demonstrated the most promising results in terms of structured voluntary activities that can cultivate positive adolescent growth. Antisocial behaviour (Wilson & Lipsey, 2000) and general conduct, social acceptance (Garst et al., 2001), negative affect (Fry & Heubeck, 1998), group cohesion (Ewert & Heywood, 1991; Glass & Benshoff, 2002), physical and expressive courage (Whittington & Mack, 2010), psychological empowerment (Shellman & Ewert, 2010), leadership, and academic achievement (Hattie et al., 1997) reflect a sample of the wide variety of outcomes that have been positively influenced by adventure-based programming. Positive effects have also
been evidenced in programmes specifically targeting adults (Hattie et al., 1997), college students (Bobilya & Akey, 2002), adolescent girls (Whittington & Mack, 2010), low achieving high school males (Marsh & Richards, 1988), high achieving adolescent males (Graham & Robinson, 2007), and juvenile offenders (Walsh & Russell, 2009; Wilson & Lipsey, 2000). The diversity of outcomes and populations that have been affected by adventure-based programming attests to its utility as a positive development strategy.

A few studies should be highlighted for their particularly remarkable effects. Marsh, Richards, and Barnes (1986a; 1986b) showed that a significant positive effect of adventure programmes on multiple facets of self-concept can be sustained 18 months following programme completion. Hattie et al.’s (1997) widely cited meta-analysis of adventure programmes demonstrated that a significant effect size of .34 found immediately following the programme was not only sustained but benefits continued to accrue after the programme and the long-term overall effect of .51 was greater than those typically found for other types of educational interventions. In addition to this, Neill & Dias (2001) found the average change in the resilience scores of Australian OB participants after involvement in the courses to be three times the effect size of the average post-programme effect found by Hattie et al. (1997). Adventure programmes have also been found to influence personality and developmental outcomes that are generally difficult or slow to change, for example locus of control (Hans, 2000) and growth across psychosocial development stages (Duerden, Widmer, Taniguchi & McCoy, 2009).

The above results would suggest that there are some inherent properties consistent across the majority of adventure-based programmes that promote positive growth. On the other hand, as one may expect, the above results cannot be generalised to all programmes. Effects are not always sustained. Garst et al. (2001) found that the increase in social acceptance in a sample of youth from a city in the Southwestern U.S. after a three day adventure experience trip was not evident four months later. Similarly, increases in social skills detected immediately following
participation in a New Hampshire 4 H summer camp had decreased significantly a few months later (Shirilla, 2009).

In other instances, no differences have been obtained at all. A comparison of 41 college students in a three month long non-residential outdoor adventure education curriculum to 29 students in a travel and tourism course did not generate any significant differences between the groups on a number of positive psychology constructs including mental toughness, hardiness, dispositional optimism, self-esteem, self-efficacy, and emotional affect (Sheard & Golby, 2006).

Unfortunately, unintended negative effects have not been completely absent from the literature. There have been instances where participants concluded programmes with worse outcomes than they began with. Orren and West (2007), for example, assessed the impact of a few brief (1 to 3 day) wilderness programmes on a diverse group of participants (African American, Asian American, Caucasian, Hispanic, and bi-racial or other). The African American participants experienced a decrease in self-concept over time. Another disappointing result is the decrease in self-concept that a group of Japanese students reported after travelling to Australia to participate in an OB programme (Purdie & Neill, 1999).

**Participant Characteristics**

In the study on social support and resilience in Australian OB programmes conducted by Neill and Dias (2001), they demonstrated that one individual’s supportive behaviour and attitude can impact the degree to which overall group benefits are experienced. Walsh and Golins (1976) also asserted that the success of the Outward Bound Process is conditional on having motivated learners from the outset. Evidence of an indirect association between programme participants’ motivation and expectations to learn and change and later personal development has surfaced. These individual differences are linked to participants’ experiences during the programme which can affect later outcomes (Sibthorp, 2003a; Sibthorp & Arthur-Banning, 2004). The greater effect sizes for adult versus adolescent adventure education groups found in Hattie et al.’s (1997) meta-analysis was also interpreted as probable differences in individual motivation.
The question of whether adventure programmes are differentially effective for males and females has been addressed by several researchers but results have not been conclusive (McKenzie, 2000; Sibthorp, 2003a). McKenzie’s (2000) review of how adventure programme outcomes are achieved described a study which demonstrated that males and females place different value on different programme activities, but whether this translates to differences is debatable. In her investigation of the experiences of participants in OB programmes in Western Canada, females reported a greater number of benefits (McKenzie, 2003). Caines and Center (2010) also found that females exhibited better outcomes than males following participation in an outdoor programme for youth. Ewert and Heywood (1991) found that males and females had similar expectations and, other than females being more dependent on their group leader, levels of interdependence, conflict and cohesion were similar. Glass and Benshoff (2002) found no differences in group cohesion between males and females after participating in a low elements challenge course and although, many of the studies in Hattie et al.’s (1997) meta-analysis did not report separate effects for each gender, the results that were available suggested that there are generally no differential effects.

Another question that should be raised is whether a participant’s ethnic background can have a significant effect on the resulting outcomes. An investigation of three different Singapore-based OB programmes revealed long-term benefits for participants (Gassner & Russell, 2008) and Glass and Benshoff (2002) obtained no significant differences across ethnic groups (Caucasian, African American and Hispanic) in the degree to which group cohesion developed from the beginning to the end of the programme. The negative effects for African Americans in Orren and West’s (2007) study and for the Japanese students in Purdie and Neill’s (1999) research were explained as a cultural mismatch. That cultural factors may have been at play in generating these effects begs the question of how adventure programming fits with New Zealand culture.
**ADVENTURE PROGRAMMES, NATURAL LANDSCAPES, AND THE NEW ZEALANDER MENTALITY**

The multifaceted topography of New Zealand makes it a land built for outdoor adventure. Within the 268,680 square kilometres (km) that make up the North and South Islands of the country (Statistics New Zealand, 2011) one has the opportunity to explore mountains, volcanoes, glaciers, caves, fjords, and native bush ([www.tourism.net.nz](http://www.tourism.net.nz)). Beyond all of this, there is more than 15,000 km of coastline and the ocean can be reached within 130km of any inland location (Statistics New Zealand, 2011).

The notion of character development through challenging encounters with nature may have particular resonance for New Zealanders not only because of the readily available opportunities for outdoor adventure awaiting them in each region of the country but also because of notable adventure role models such as Sir Edmund Hillary. When Sir Edmund Hillary reached the top of the highest peak in the world in 1953 with Sherpa Tenzing Norgay (Calder, 2008; Sir Hillary, 1999) he demonstrated the heights that could be reached by any ordinary person. His attitude of perseverance through challenges (both physical and mental) and humility in one’s accomplishments has been imbued in the national psyche, as evidenced in the claim below.

A survey in 1996 named the conqueror of Everest as the top choice of men and women of all ages and across all ethnic groups – far ahead of any All Black, any political or artistic figure – as the person who best embodied “the spirit and essence” of our small nation. (Calder, 2008, paragraph 2).

The momentum for the outdoor education movement in New Zealand began in the 1930s and 1940s with primary school camping trips, much in the same way as it did in Britain (Lynch, 2006) and in the U.S. (Ewert & Garvey, 2007). In the 1950s, outdoor camps were allowed to go ahead during school time. Outward Bound New Zealand (OBNZ) got underway in 1962 ([www.outwardbound.co.nz](http://www.outwardbound.co.nz)) which subsequently provoked an increase in adventure-based activities as a main means of outdoor education for secondary school students (Lynch, 2006). Nevertheless, it was not until 1999 that outdoor education became a formal part of the compulsory educational curriculum (Lynch, 2006).
According to Lynch (2006), the growth of outdoor education (and especially its adventure-based derivative) has encountered several roadblocks over the course of its 80-plus year existence in New Zealand society. There has always been a continuous struggle to obtain adequate funds and this has affected the quality of instructor training as well as the number of qualified staff, which unfortunately led to some devastating tragedies. Dissident opinions have been expressed over the frivolous expenses incurred by adventure activities during hard economic times but, in general, it has been sustained community advocacy that has kept the movement afloat through some tumultuous periods. In the late 1970s, five fatalities occurred as a result of outdoor education activities, but the fact that community and even family members of those who died responded with calls for improved resources rather than a shutdown of operations provides a clear illustration of the public belief in the potential benefits of outdoor education.

**Is the Belief Justified?**

A review of New Zealand-based evaluation research on adventure-based programming would seem to support the overall positive intra and interpersonal effects found in similar initiatives overseas. Within the school context, students involved in curriculum-based outdoor pursuits have attested to the idea that outdoor education experiences promote new coping skills (Lynch, 2006), a positive view of self and the future (Davidson, 2001), and an increase in social connectedness (Davidson, 2001; Lynch, 2006; Smith, et al., 2010).

In terms of OBNZ courses, Martin and Leberman (2005) found an impressive overall effect size of .68 across four different courses (accommodating both youth and adult participants) when assessing changes from pre to post programme in a composite life effectiveness measure. A caution to be prudent when interpreting the effect due to the small samples of each programme notwithstanding (n = 13 to 56), the positive outcomes were corroborated with comments made in follow-up interviews. Participants reported gains in self-awareness, self-confidence, perseverance, and respect for others.
Martin and Legg (2002) also used a mixed method approach to compare the results from 54 participants in standard 22-day OBNZ courses to those of 39 adults in the shortened 9-day versions. Positive gains in three of the 17 self-concept subscales amounted to a significant overall positive effect of .32 for the standard courses but the shorter courses did not yield much change. Qualitative responses to open-ended questions six months following the courses indicated that increased self-confidence was the primary outcome whilst over 70% of the participants from both versions of the programme also claimed to have grown interpersonally, mainly with regards to having better relationships.

Martin and Legg’s (2002) finding that the participants in the shorter 9-day courses reported no significant improvement immediately following their OBNZ courses point to the moderating effect that programme duration can have on programme success, which is in line with other overseas studies (Garst et al., 2001; Hans, 2000; Hattie et al., 1997). Other New Zealand-based studies reporting on optimal programmatic elements and conditions also align closely with the Walsh and Golins (1976) model and studies conducted overseas, as described above. The novel physical and social environment (Lynch, 2006; Smith et al., 2010), the unpredictable but peaceful atmosphere (Lynch, 2006), the more engaging informal structure (Lynch, 2006; Smith et al., 2010), the element of choice or autonomy (Davidson, 2001; Lynch, 2006), the challenging activities (Davidson, 2001; Eggleston, 2000), and an empathetic and supportive group and instructor (Eggleston, 2000; Martin & Leberman, 2005) are all factors that New Zealand participants have voiced as being meaningful to their growth in adventure experiences.

**Is it a fit for all New Zealanders?**

As articulated in the YDSA, an inclusive society of flourishing young New Zealanders is contingent upon the adherence to bicultural principles; hence, it is important to assess whether the general principles of adventure programming described above align with Māori tikanga (values-based protocols/guidelines). Cormack (1997) has delineated the optimal learning environment for young Māori.
Fitting in with the dominant Pākehā culture while staying true to the interdependence customarily found in Māori practices, necessitates the incorporation of group, pair and individual activities. The instructor should gradually transfer responsibility to the group then to individual students after setting boundaries for the collective. Cormack encourages educational approaches that capitalise on the full range of human senses and provide learning opportunities that are grounded in reality so as to ensure relevance to the learner.

Viewing the field of outdoor education in Australia from a critical lens, Lugg (2004) suggests that the current outdoor programming strategies are a by-product of a colonialist ethos that perpetuates an individualistic and masculine portrait of personal growth – the “making men out of boys” (p. 6) mentality. Early programmes were indeed off-limits to women (Hattie et al., 1997; Lugg, 2004) and while self-sufficiency is encouraged (Lugg, 2004), it is arguable that the underlying values of current programmes promote individualism or a competitive spirit. Programmes now cater to mixed gender and all female groups and are often structured in a way that also promotes the values of care and community (Quay et al., 2000). Walsh and Golins (1976) described the social environment of the Outward Bound Process as “individuality within a cooperative framework” (p. 6) which recapitulates the philosophy of advocating individual decisions only if they are supported by the group and account for collective desires and well-being. This social structure and the experiential learning criteria that are undeniably the crux of adventure programming are remarkably similar to the classroom conditions Cormack (1997) suggested would be favourable for Māori achievement.

A study of the month-long Whakapakari programme on Great Barrier Island, New Zealand for Māori adolescents with backgrounds of neglect, abuse and high risk behaviour provides some insight into the suitability of outdoor education programmes for this demographic. Supporting the notion of outdoor programmes as an avenue to community mindedness and an ethic of care (Quay et al, 2001), the interviews that took place 18 months after the programme with ten participants indicated that it was the “whānau” atmosphere that was the most salient aspect of the experience and it produced flow-on benefits for their
relationships outside of the programme (Eggleston, 2000). Some remarked that the discipline, the routine, and the high level of effort needed to accomplish tasks helped them to develop and the majority of the participants (n=6) indicated that their current interest in Māori culture was seeded in the Whakapakari programme (Eggleston, 2000). It is necessary to point out that Māori culture was explicitly embedded in the programme with opportunities to learn indigenous cultural practices.

For many of the participants, the dissimilarity between the programme environment and their home lives prevented adequate transference of the lessons learned. Some of the values they developed within the Whakapakari whānau were simply not adaptive for their day to day environment because reciprocity was lacking (Eggleston, 2000). Although successful transference may be the mark of a successful programme (Gass, 1999), it cannot be expected to produce discernible effects in the long-term if the values and practices are not supported by the encompassing social systems. If forces within the socio-ecological system are working against the progress made, is it appropriate to designate the programme as a failure?

In any case, one cannot wholeheartedly accept the results of the Whakapakari programme as representative of other programmes serving Māori youth for two main reasons. First, while it was set in the wilderness, it was a wilderness therapy programme not a wilderness adventure programme. High adventure pursuits were not part of the curriculum. Second, the participants were already manifesting problem behaviours prior to participation and came from backgrounds of neglect and/or abuse. These are not invariable characteristics for the general Māori youth population. In fact, Māori tikanga underscores the values of manaakitanga (hospitality and kindness) and aroha (concern, compassion, and love).

With regards to the Pasifika and Asian youth sectors, the communal and interdependent nature of the adventure programme context may also have particular appeal for participants from these collectivist-based cultures. Two of the
four participants interviewed by Davidson (2001) after involvement in a school-based outdoor education curriculum study were originally from a Pacific island nation and both of these individuals described the course as a positive experience, but these experiences can hardly be generalised to others sharing their ethnicity. The unfortunate outcome for the Japanese students in the Australian OB programme (Purdie & Neill, 1999) raises legitimate concerns about the suitability of these programmes for Asian youth in the New Zealand context; however, the situation may be quite different for youth exhibiting higher degrees of acculturation.

The cultural shifts that have occurred in New Zealand society over the past decade brings up a number of interesting questions for adventure programming researchers and practitioners. The growing Asian and Pacific populations make the question of cross-cultural suitability more pressing. The introduction of compulsory outdoor education as part of the formal education curriculum in 1999 (Lynch, 2006) makes one wonder whether or not this will have an desensitising effect on the excitement and motivation of participants involved in commercial or not-for-profit programmes. Will outdoor education leaders still be able to rely on the unfamiliarity of the situation to act as an important catalyst for change? Brett (as cited in Martin & Legg, 2002) noted a decline in registrations of OBNZ participants in the late 1990s. He thought this could be due in part to the negative publicity following two fatalities in the early nineties and his belief that New Zealand society appears to have become more risk-averse in recent times (Davidson, 2004). Will the belief New Zealanders have in the benefits of outdoor adventure programmes prevail as it has in the past, even after the most recent adventure programme canyoning tragedy at Mangetopopo gorge (Binning, 2000)? These unanswered questions suggest that continued research on the processes and value of New Zealand-based adventure programmes is needed.
DEVELOPMENT THROUGH SERVICE

“How can we expect our youth to become good citizens, if we don’t provide them with opportunities to care, to participate, and to practice as members of a democratic society?” (Terry & Bohnenberger, 2003, p. 30)

Many argue that the past few decades have been characterised by a deterioration of the civic spirit because people today are less likely to vote, join political parties, have membership in other types of organizations (Carpini, 2000; Putnam, 1996; 2001; Vowles, 2004), be informed about the world and political issues, and they have less trust in others in general (Carpini, 2000; Putnam, 1996; 2001). This downturn in civic engagement is even more pronounced for individuals in younger cohorts (Blais, Gidengil, Nevitte & Nadeau, 2004; Carpini, 2000; Gallego, 2009).

In New Zealand the picture looks slightly better. Barr (1998) remarked that New Zealand has a respectable performance of civic responsibility on the international scale: newspapers readership is high; the evening news receives the highest viewer ratings; and the high levels of library memberships and book ownership point to a well-informed population. Beyond this, voter turnout in New Zealand has been high relative to other nations (Vowles, 2004). The 2010 Social Report produced by the Ministry of Social Development indicated that, in 2006, New Zealand was the sixth most socially trusting nation of 25 OECD countries (Ministry of Social Development, 2010). Results from the Youth ’07 secondary school survey suggests that social trust may be even higher in young people; 83% of the sample reported that they felt their neighbours were trustworthy (Adolescent Health Research Group, 2008), whereas social trust levels in the general population sat around 76% in 2006. This is encouraging as Putnam (2001) has proclaimed that social trust is the bedrock of a democratic society.
THE ALIENATION OF YOUNG PEOPLE FROM COMMUNITY LIFE

All the same, volunteer rates among secondary school-age youth in New Zealand appear to be quite low. According to the Youth ’07 finding, only 14% of the sample reported having helped others in the community in the past 12 months and only 4% indicated that they belonged to a volunteer organisation (Adolescent Health Research Group, 2008). Moreover, a study by Panelli, Nairn, Atwool, & McCormack (2002) demonstrated that media portrayals of New Zealand youth in the “public space” (p. 38) are abundantly negative. By investigating the descriptions of young people and their occupation of public spaces in newspaper articles from the Otago Daily Times over the course of nearly one year (2001-2002), they showed that the idea of young New Zealanders occupying public space is linked to discourses of threat to others and, consequently, exclusion from or increased control of their access to these spaces.

These kinds of discourses and the fact that young people are rarely afforded the opportunity to counter negative media claims (see Crane, n.d) send a clear message that young people are not valued as full citizens (Panelli et al., 2002; Youth Affairs Council of Victoria, n.d.). Furthermore, it reflects a common divide between young people and adults in contemporary Western society (Youniss, McLellan, Su & Yates, 1999). When viewed from this vantage point the deterioration of civic engagement in the younger demographic is not so surprising.

Strategies to Change the Profile of Youth in New Zealand

The Ministry of Youth Development (MYD) is committed to forging opportunities for youth participation in the New Zealand government and creating space for young people to voice their ideas and concerns. One MYD initiative that is coordinated by Ara Taiohi, an organisation that supports people and organisations that work with youth is Youth Week (Ministry of Youth Development, 2011). Youth Week aims to put young New Zealanders’ achievements to the forefront and celebrate their strengths and diversity in hopes that this will promote a society that values their young people. MYD also

---

8 The Otago Daily Times is a newspaper based in Dunedin, New Zealand.
continually advocates for youth input into government decisions and in broader community affairs and regularly consults with youth advisory groups when developing relevant policies (www.myd.govt.nz) but the above section suggests that more could be done at the grass-roots level.

**SERVICE-LEARNING AS A SOLUTION**

Overseas, youth service-learning programmes are seen to offer an avenue towards the rectification of community-youth ties (Terry & Bohnenberger, 2003; Youniss et al., 2002). Most often found as an integrated part of or as an adjunct to an academic curriculum, service-learning programmes provide students with authentic learning experiences that address a real community or societal need (Billig, 2000; Furco, 1996; Scales, Blyth, Berkas, & Kielsmeier, 2000), for instance volunteering for a community-based organisation, developing a project to improve the environment, or tutoring younger children (Billig, 2004). As a result, young people and the wider community obtain bi-directional rewards (Furco, 1996) – service and learning. Service-learning shares a common lineage with outdoor education programmes to the experiential education movement (Billig, 2000; Conrad & Hedin, 1991; Eyler, 2002; Prouty, 2007; Seaman & Gass, 2004). John Dewey was a strong advocate of democracy and while the broad message of his educational philosophy may have been that concrete interactions with the environment provide educative experiences (Baker, 1966; Dewey, 1963) he also saw education as a means to develop democratic citizens (Baker, 1966; Youniss et al., 2002). His idea of “education as democratic socialization” (Seaman & Gass, 2004, p. 72) was picked up through service-learning programmes.

**Considerations for Service-Learning Programmes in New Zealand**

“Service-learning” is an American construct that became popular in the 1980s and is based on the ingrained American ethos of volunteerism. While other nations took these ideas on board in the 1990s because the main tenants of civic engagement were applicable, other elements may not be easily transposed onto other cultures (Iverson & Espenschied-Reilly, 2010). Irish service-learning practitioners, for example, pointed to the fact that the word “service” is couched in
punitive or subservient connotations in their culture (Iverson & Espenschied-Reilly, 2010).

Service-learning programmes exist here in New Zealand, at least in the tertiary education sector (Perry, 2009), but the term is not usually found in the common vernacular and research on New Zealand-based initiatives is difficult to find. One may assume that implementing service-learning programmes in the New Zealand context could only reinforce the positive messages put forward by MYD and Ara Taiohi and promote closer community ties, which is important for positive youth development (Benson, 1997; Youniss et al., 1999); however, it is important to first reflect on the fit of these programmes for New Zealand society, as I do below.

In New Zealand, the idea of “volunteering” in a formal sense stems from a dominant status model. According to Wilson (2001), the term “is frequently associated with formal social service volunteering and the ‘middle-class, middle-age’ stereotype.” (p. 2). Yet, activities that would clearly be categorized as voluntary from this perspective may not be identified as such by people from different socioeconomic brackets or those who do not form the ethnic majority (Wilson, 2001). Māori and Pacific notions of family and community (Wilson, 2001) and the implied obligations Māori have to reciprocally contribute to their whānau, hapu, and iwi also complicate the connotations of the term (Robinson & Williams, 2001). An example of this cultural disconnect is evidenced in the following quote:

When I get up as a Pakeha and mow my lawns, I mow my lawns...When I go down the road to the disabled children’s home and mow their lawns I volunteer to do something for the other...When my friend Huhana gets up and mows her lawns, she mows her lawns, when she goes down to the Kohanga Reo and mows lawns, she mows her lawns. When she moves across and mows lawns at the Marae and the Hauora, she mows her lawns – because there is no sense of ‘other’. (Stansfield, 2001 as cited in Wilson, 2001, p. 22)

The rates of volunteering by Māori appear high if accounting for informal forms of volunteering but one can see how the different beliefs of the various cultural groups that make up New Zealand can cloud precise estimates of volunteer activity (Wilson, 2001). This is not to say that confusion over terminology should preclude the implementation of service-learning programmes in New Zealand but as with any imported programme model, care should be taken to
ensure that the language, values and objectives of the programme are modified to match those of the recipient culture. Project K has made the community service component of the programme their own by reframing it as a “Community Challenge” and by moving away from the traditional curriculum-integrated model of service-learning to suit their own needs and objectives, putting their programme in good stead in this regard. These slight modifications aside, many features of service-learning are found in Project K’s Community Challenge thus it is helpful to review the important elements, influencing factors and evidence of effectiveness associated with service-learning programmes. Hence, I address these below.

THE CENTRAL ELEMENTS & OPTIMAL CONDITIONS

Researchers and practitioners across this discipline have not yet come together on a solid theoretical foundation for service-learning (Seaman & Gass, 2004); however, several researchers have put forth ideas of the essential components. Four elements that are consistently mentioned are: reflection, youth voice and agency, and meaningful activities that meet real community needs (Billig, 2000; Dymond, Renzaglia & Chun, 2007; National Clearing House, 2004). Notably these are reminiscent of the factors that are thought to underlie growth in adventure programmes. Four interrelated factors which also potentially influence the quality of service-learning programmes include good relationships with community partners, careful planning, dosage and congruence between activities and goals.

Reflection

Eyler (2002) submits that the amount and kind of reflective practices employed are the most critical dimensions of programme quality. Reflection, which equates to the way in which participants process and make meaning of their service-learning experiences, is the primary vehicle of learning in this context (Seaman & Gass, 2004). Eyler (2002) urges practitioners to purposefully integrate reflective activities before, during, and after the service project(s) and expresses concern over those who make a feigned effort to simply include reflective activities as a tokenistic add-on to the curriculum. Other researchers have expressed a
similar sentiment. Conrad and Hedin (1991) indicated that, based on their experiences, the inclusion of a reflective seminar noticeably influenced programme success. Rutter and Newmann (1988) claimed that their most important recommendation was ensuring that a reflective seminar that enabled participants to critically reflect on social responsibilities was included and that all project plans accounted for the time it takes to engage in in-depth reflection. Furthermore, comparisons of programmes with and without explicit reflection components have revealed that these components can intensify positive effects (Scales, et al., 2000).

In contrast, Waldstein and Reiher (2001) discovered that, for their sample of 801 ninth grade participants from six different schools across three U.S. states, individuals gained more personal development from service-learning experiences without a definitive reflective component than those in programmes with one. They concluded that perhaps participants in the “no reflection” programmes were engaging in informal reflection on their own or it may have been the way in which reflection was approached that detracted from the experience. Student ownership may be compromised if reflective activities are dictated from above and make use of strategies that are akin to homework activities (Waldstein & Reiher, 2001). This further highlights that the manner in which reflection is approached is as important as its inclusion in the programme.

Participant Voice and Agency

Second only to reflection, participant voice and agency is another priority topic in the service-learning literature. In addition to a lack of voice in general public affairs, even explicit attempts to include young people in organisational processes often become tokenistic or even manipulative gestures (Dallago, Cristini, Perkins, Nation & Santinello, 2010; Evans, 2007; Fredericks, Kaplan & Zeisler, n.d.). In his investigation of sense of community in adolescents involved in two youth development programmes based in Nashville, Tennessee, Evans (2007) discovered that the majority of the sample suffered from a deficiency in their sense of being able to influence their communities. However, in instances where youth were genuinely invited to express their opinions they were excited by the
opportunity to get involved and their participation reinforced a desire to contribute to their communities.

Along these lines, interviews with students from eight U.S. high schools with service-learning programmes indicated that it was the increased opportunities for agency and ownership and being afforded the same level of respect as adults that distinguished their community service fieldwork experiences from the traditional classroom lessons, and it was the fieldwork experiences that generated more favourable reviews (Rutter & Newmann, 1989). Billig (as cited in Fredericks et al., n.d.) asserts that the youth voice concept is not only integral to service-learning programmes but it has the ability to truly maximise any positive effects. Due to this, Fredericks et al. (n.d.) claim it is “one of the hallmarks of a successful program” (p. 2). This belief was echoed by the service-learning practitioners that Dymond and her colleagues (2007) consulted and by members of the National Service-Learning Cooperative (National Youth Leadership Council, 2004).

Youth voice and agency is directly linked to the type of support provided by the adults involved. Beyond having access to influential members of the community, young people need support to develop the knowledge and skills to meaningfully participate in these opportunities (Evans, 2007). Nelson and Eckstein (2008) suggested that this requires providing participants with a general framework but allowing the young people to fill in the details and do the necessary work. There is a fine line between providing enough support to keep the activities on track while allowing students to take ownership of the process. One objective may be more important than the other depending on the goal of the programme. Nevertheless, it appears that the ability to maintain this balancing-act is a feature of an exceptional programme coordinator (Larson, Walker, & Pearce, 2005).

**Authenticity and Community Links**

All types of service-learning activities are not equally effective when it comes to generating the desired outcomes. In a broad sense, the specifics do not matter much but the authenticity of the activities do, as several researchers have
pointed out (Billig, 2000; Dymond et al., 2007; National Youth Leadership Council, 2004). Many believe this equates to conducting projects that are genuinely valued by the community in question (Billig, 2000; McBride, Pritzker, Daftary & Tang, 2007; Rutter & Newmann, 1989; Terry & Bohnenberger, 2004) and others believe that meaningful experiences also require face to face interactions with the recipients of the service (McBride et al., 2007). For these reasons, careful planning and having strong connections with community agencies should be a prerequisite to community-building projects. Having strong partnerships with community agencies to begin with may increase the responsivity of those receiving the services to the young people providing them (Dymond et al., 2007; National Youth Leadership Council, 2004).

**Dosage**

Unfortunately, not having enough time is a common barrier to the inclusion of service-learning in an academic curriculum but meaningful service experiences require adequate dosage (intensity and duration). Research by Peter Scales and his associates showed that individuals with a few weeks or more of service experiences had a higher number of the Commitment to Learning assets (a category of the Developmental Assets® framework outlined in the first section of this chapter) than both those who had no service learning and those with only a few hours (Scales, Benson, Roehlkepartain, Sesman & van Dulman, 2006; Scales & Roehlkepartain, 2005). Scales et al. (2000) also found that middle school students across three U.S. states who had accumulated more than 30 hours of service-learning experiences had more self-efficacy for helping others than students with fewer hours or no experience at all. Interviews with university-sponsored AmeriCorps volunteers indicated that the length of service experiences matter (Einfeld & Collins, 2008). These volunteers felt that longer services experiences allow more time to build connections with community members and, for them, this engendered greater multicultural competence.
Congruence with Goals

Ensuring the above areas are attended to should facilitate success but in the end any positive gains will be difficult to demonstrate if programme activities do not align with the stated goals. Specifics do matter when it comes to quantitative evaluation of the programme. If one aspires to detect changes in specific outcome constructs then there needs to be congruence between the programme structure and activities and the intended goals (Billig, 2000; Einfeld & Collins, 2008; Rutter & Newmann, 1989), whether these are academic, career, personal, social, or civic development.

THE THEORETICAL PROCESSES OF SERVICE-LEARNING

While an understanding of the main components and optimal conditions for service-learning is informative, Terry and Bohnenberger’s (2003; 2004) work provides a more intricate picture of the underlying mechanisms of change. Their blueprint for K-12 service-learning accommodates varying degrees of service-learning while distinguishing between the more service-oriented, learning-oriented, and equally weighted service/learning approaches by placing each along a service-learning continuum. Each of the Community Service, Community Exploration, and Community Action approaches is associated with graded levels of reflection which require increasingly complex cognitive processes.

Their description of this framework fleshes out some of the likely processes thought to be operating within each approach. For the Community Service approach, when the balance is tipped towards service, reflection occurs at the observational level and participants acquire knowledge through simple problem solving and communication. Personal growth via increased efficacy and esteem, moral growth (for example altruistic values and group responsibility), and an increase in school engagement are some of the outcomes that may result. The balance is weighted towards learning during the Community Exploration approach and the levels of service will vary depending on the activities. Reflection becomes more analytical and knowledge is applied rather than simply acquired.
The nature of the activities will also dictate the types of skills that are developed. In addition to the previously cited outcomes, awareness of oneself and of the larger society generally grows and a vocational identity begins to emerge.

The highest levels of service and learning culminate at the Community Action level and Terry and Bohnenberger (2003) provide an extended explanation of this approach. The processes at this level can be conceptualised as a mixture of Brown, Collin’s and Duguid’s (1989) cognitive apprenticeship model, the Osborne-Parnes Creative Problem Solving Process, and cooperative learning strategies. To clarify, the cognitive apprenticeship model uses a graded mastery approach to teaching whereby the instructor moves from modeling the required skills, to coaching the students in employing the skills themselves, and finally fading into the background to allow the student to exercise the skills independently (Brown, Collin’s & Duguid, 1989; Terry & Bohnenberger, 2003). The Creative Problem Solving Process follows a sequence of steps from identifying a challenge, articulating the issue, providing alternative strategies to deal with the problem, evaluating the alternatives and enacting the plan. The instructor supports the students’ autonomy by guiding the process but allowing them to choose the content and direction of the task at hand (Terry & Bohnenberger, 2003). Cooperative learning strategies utilize action-oriented subgroups working independently on tasks that are then coordinated in order to move the entire collective towards a common goal (Terry & Bohnenberger, 2003). An explicit focus on reflection features throughout the service project and, at this level of reflective synthesis, knowledge is integrated and multiple viewpoints of a situation can be contemplated. In concurrence with Billig (2000) and Dymond et al. (2007), a concluding celebratory component to acknowledge and reinforce the students’ accomplishments is recommended (Terry & Bohnenberger, 2003).

**EVIDENCE OF EFFECTIVENESS IN VIEW OF POPULAR SUPPORT**

Rutter and Newmann (1989) estimated the proportion of U.S. high-schools offering service-learning programmes to be at about 27% in 1984 and, according to a U.S. Department of Education survey, by 1999 percentages had risen to around
83% (Fredericks et al., n.d.; Kielsmeier, 2000). Kielsmeier (2000) equates these percentages to an increase in participation from 81,000 high school students in 1984 to 2.9 million 15 years later. Given this explosion in popularity, it is reasonable to review whether these programmes are effective in producing meaningful outcomes.

While the long-term agenda of these programmes (including Project K) may be to influence civic engagement at a societal level, the effects of these endeavours on community-level outcomes are rarely investigated (Conrad & Hedin, 1991). The few studies that have been conducted have provided support for the programmes’ capacity to change public perceptions of youth (Billig, 2000, 2004); however, as community-level data were not available for this thesis, the focus of this research necessarily investigates effects at the individual level.

Theoretically speaking, youth service-learning experiences should help to cultivate a civic identity because opportunities are provided to engage with individuals who have different worldviews, whilst the activities require adherence to norms that allow young people to move beyond their differences in order to achieve collective goals (Yates, & Youniss, 1998; Youniss, McLellan & Yates, 1997). Service activities also cultivate developmental needs such as agency, social connectedness and moral-political awareness which are important precursors to a civic identity (Yates & Youniss, 1996).

Benefits such as these are likely to arise because service experiences address a number of assets within the Developmental Assets® framework (Benson, 1997; Nelson & Eckstein, 2008; Scales et al., 2000). Forging connections with supportive adult role models who provide boundaries and high expectations, empowering conditions where young people’s opinions are valued and social competencies are developed, and the encouragement of positive values from activities that are a constructive use of their time are all important features of the Developmental Assets® framework and all can be experienced during a service-learning project (Benson, 1997). Accordingly, personal, social and civic domains are common

---

9 Billig, n.d. clarifies that this proportion varies greatly depending on how broadly service-learning is defined. Regardless, the surge in programmes is undeniable.
outcome categories found in the service-learning literature, however, academic or cognitive and vocational outcomes are also relevant (Billig, 2000; 2004; Nelson & Eckstein, 2008).

Considering the rationale for programmes is based on a need for civic engagement, it is encouraging that some results have been demonstrated in the civic domain. An increased awareness of community and political issues (Billig, 2004), a greater sense of civic responsibility (Billig, 2004; Dallago et al., 2007; Waldstein & Reiher, 2001), adherence to conventional norms (Youniss et al., 1999), and future intentions to be involved in civically-oriented activities (Hamilton & Fenzel, 1988; Lakin & Mahoney, 2006; Waldstein & Reiher, 2001) have all been evidenced in empirical studies of service experiences. On the other hand, the results are not overwhelmingly positive (Billig, 2004; Conrad & Hedin, 1991; Rutter & Newmann, 1989). For example, across eight U.S. high school service-learning programmes, Rutter and Newmann (1989) found only negligible effects on civic responsibility. It should be noted, however, that the solidification of a civic or political identity may take time to develop and a few studies have shown promise that the effects of service experiences in adolescence are linked to features of a long-term civic identity (Beane, Turner, Jones & Lipka, 1981; Billig, 2000; Yates & Youniss, 1996; 1998).

The majority of evidence in favour of service-learning programmes has been in the personal and social development arenas (Billig, 2004; Rutter & Newmann, 1989). Researchers have uncovered positive outcomes for self-esteem, self-efficacy, moral development (Billig, 2004; Conrad & Hedin, 1991; Yates & Youniss, 1996), empathy (Conrad & Hedin, 1991; Lakin & Mahoney, 2006), concern for others (Scales et al., 2000), social responsibility, (Billig, 2000; 2004; Conrad & Hedin, 1991) helping behaviours, awareness and tolerance of diversity (Billig, 2000; 2004; Yates & Youniss, 1996), trust, and better relationships (Billig, 2000; Yates & Youniss, 1996).

Fewer studies have addressed academic and career-related outcomes and although the evidence is not as strong as that for personal or social characteristics
Adventure, Service & Mentoring for Youth

(Billig, 2004; Conrad & Hedin, 1991), the general pattern has been one of positive effects (Billig, 2000, 2004; Conrad & Hedin, 1991). For example, Shelley Billig’s reviews of the literature revealed an optimistic picture for academic achievement, school engagement (Billig, 2000, 2004), classroom behaviour, homework completion, drop-out rates and attendance (Billig, 2000). In terms of the vocational domain, Billig reported studies showing that students with experiences of service had stronger job skills (Billig, 2004), better attitudes towards work (Billig, 2000), were more aware and knowledgeable about career options (Billig, 2000, 2004) as well as more realistic (Billig, 2000). Additionally, Johnson, Beebe, Mortimer, & Snyder (1998) found that participation in volunteer activities reported by individuals in a sample of nearly 1000 ninth grade students from Minnesota predicted an increase in intrinsic work values over time. Participation also predicted a decrease in the anticipated importance of a career but the researchers believed this may have been due to the exposure to a greater variety of options and the focus on the intrinsic benefits of work.

The potential for adverse effects exists with any kind of social programme when quality is not controlled and service-learning is no different. For example, an Education Commission of the States (Fredericks, n.d) report warned that unless young people are genuinely involved in the process, service-learning experiences can further reinforce the idea that they are not a valuable part of society. This caveat notwithstanding, actual evidence of iatrogenic effects is scarce.

Overall, the effects have generally been tilted towards the benefits across the five main goal domains. However, mixed effects of several positive and some null effects are ubiquitous across studies (Conrad & Hedin, 1991; Eyler, 2002; Waldstein & Reiher, 2001). Furthermore, previous reviews have pointed out that many of the positive effects are small (Eyler, 2002). This is because the way in which the environmental, programmatic, and participant factors come together is the true determinant of programme success. The less than impressive impact of service-learning on outcomes across studies has been attributed in large part to the enormous diversity found across service-learning activities (Furco, 1996; Hamilton & Fenzel, 1988; McLellan & Youniss, 2003) and the implementation of these
activities (Eyler, 2002). For example, one may encounter “activities that range from the Saturday afternoon car wash for charity, to tutoring peers in one’s classroom, to regular stints at a soup kitchen” (McLellan & Youniss, 2003, p. 48) within the same service-learning literature base. Other factors that have been acknowledged to potentially affect the strength of the results include the diversity in the responses of different individuals to similar activities (Hamilton & Fenzel, 1988); thus I turn to these next.

**PARTICIPANT-ENVIRONMENT INTERACTIONS**

With regards to research on participant characteristics and service-learning, most of the emphasis has been on which types of individuals volunteer. For example, the links between volunteers rates across both genders has been investigated (Flanagan, Bowes, Jonsonn, Csapo, & Sheblanova, 1998; Johnson et al., 1998; Vieno et al., 2007) as have associations with social background factors (Johnson et al., 1998; Yates & Youniss, 1996), personality characteristics, motivations, levels of social connectedness and personal competence (Yates & Youniss, 1996), school-related and psychological factors (Johnson et al., 1998), and risk behaviours (Vieno et al., 2007). Unfortunately, these studies do not allow us to tease out whether these factors predispose young people to want to volunteer or whether some of these factors occur as a result of volunteering nor do the data provide any indication about the differential effectiveness of service-learning programmes (which may or may not be embarked on voluntarily) for different types of people, for example across gender or socioeconomic background.

Two studies, however, provide a little insight into these factors. Hamilton and Fenzel’s (1989) research revealed that female participants in various youth service projects obtained greater benefits than males, which manifested as a higher degree of improvement in the females’ sense of social responsibility, but both sexes were equally likely to seek out further volunteering opportunities. The research by Peter Scales and his colleagues (Scales et al., 2006; Scales & Roehlkepartain, 2005) comparing schools catering to students from different socioeconomic status (SES) brackets suggested that service-learning programmes have the potential to
minimize the achievement gap typically found between low and high SES students.

In sum, service-learning is a promising avenue for cultivating positive youth-community ties but definitive conclusions about the expected impact cannot be made without additional research. Rigorous research investigating programme processes and, especially, the moderating influences of participant and wider socio-environmental influences is needed. It is the investigation of these elements that can better inform the development of similar initiatives in New Zealand including Project K.
MENTORING PROGRAMMES

John began to mentor Tyrell when Tyrell was ten years old. John took Tyrell travelling with him on several trips overseas. Tyrell developed an interest in culture and language. He graduated with honors from high school, attended a prestigious Minnesota College, and became fluent in French and German. The Mayo medical clinic selected Tyrell as one of 40 students chosen out of 4000 applicants and he became a family physician. Eighteen years after first being matched John and Tyrell are still connected. Tyrell acknowledges that his potential to fulfill his dreams of becoming a doctor may have gone untapped if it weren’t for the role of John in his life. (Pederson, Woolum, Gagne & Coleman, 2009, p. 1310-1311).

Mentoring programmes are not a contemporary phenomenon. Formal mentoring relationships have been facilitated for over a century. In 1902, the first Big Brothers agency was established, and Big Brothers/Big Sisters (BBBS) remains the largest and best-known organization promoting adult-youth partnerships in the world (Evans & Ave, 2000; Rhodes & DuBois, 2008). Approximately three million young people are thought to be currently participating in a mentoring programme in the U.S. alone and this growth does not show any signs of stopping (DuBois, Portillo, Rhodes, Silverthorn & Valentine, 2011; Rhodes & Lowe, 2009). Noticeable growth has also been noted in the UK (Philip & Hendry, 2000) and in New Zealand (Dunphy et al., 2008; Evans & Ave, 2000; Farrugia, Bullen, Dunphy, Solomon, Collins, 2010). A sound rationale for mentoring may be at the foundation of the popular support for this youth development strategy but the passion for mentoring is likely stoked by stories like that of John and Tyrell which may understandably “strike deep emotional chords” (Rhodes, 2008, p. 35) with the public.

THE RATIONALE FOR MENTORING PROGRAMMES

Access to positive adult role models and additional support from individuals outside the family is a consistent factor that separates young people that rise above adversity from those that succumb to negative lifestyles (Cowen & William, 1998; Werner, 1995). Accordingly, having non-parental adult role models to set high expectations and to provide boundaries and support is explicitly listed as assets within the Search Institute®’s Developmental Assets® framework (Leffert...
et al., 1998; Rhodes & Lowe, 2009). Scales, Benson, & Mannes (2006) provided evidence that experiences of support, empowerment and boundary setting with adults in the community is linked to reduced risk behaviour and increased thriving later in life. Corroborating the link between youth/non-parental adult bonds and resiliency, higher levels of health and well-being were evidenced in populations of youth in foster care (Ahrens, DuBois, Richardson, Fan, & Lozano, 2008) and in young African American mothers (Rhodes, Ebert & Fischer, 1992) when these young people had a non-parental adult role model in their lives.

As individuals progress through adolescence there is a natural shift to an environment that is less influenced by adults (Darling, 2005; Scales & Gibbons, 1996). Parental monitoring tends to decrease, peer relationships take centre stage, and fewer personal interactions occur with teachers (Darling, 2005). Despite this, as adolescents age some express a growing recognition of the necessity of relationships with adults at this stage in their lives (Scales & Gibbons, 1996).

Researchers have acknowledged that changes in family structures in the U.S. (i.e. the prevalence of single-parent homes and extended family members living in different geographical regions) mean that family members are spending even less time with their children (Rhodes, 2005; Thompson & Kelly-Vance, 2001). The situation is very similar in New Zealand (M. H. Durie, 1997). The Youth’07 survey revealed that only 40% of students felt they spent enough time with their fathers most of the time and approximately half felt they spent enough time with their mothers (Adolescent Health Research Group, 2008).

To make matters worse, overcrowded classrooms further reduce the time teachers are spending with individual students (Rhodes, 2005), people are spending increased time in front of the television, and adults are spending less time socialising with friends and neighbours (Putnam, 2001). Consequently, many believe there are few opportunities for young people to forge connections with potential adult role models (de Anda, 2001; Larson, 2000; Morrow & Styles, 1995; Philip & Hendry, 2000; Spencer, 2007). Just over half of the New Zealand students in the Youth ‘07 survey indicated that they have an adult in their community with
whom they can discuss serious problems (Adolescent Health Research Group, 2008). Thus it is not surprising that there continues to be widespread support for strategies that facilitate adult-youth bonds by purposefully constructing relationships that would otherwise not occur. A clear example of such a strategy is found in mentoring programmes.

**MENTORING RELATIONSHIPS**

Relationships that develop as an organic process in a young person’s normative environment fall under the rubric of *natural mentoring* and these mentors have been qualified by researchers as caring and reliable individuals other than a parent who have had a substantial impact on a young person’s life (DuBois & Silverthorn, 2005; Rhodes, Ebert & Fischer, 1992; Scales & Gibbons, 1996). A formal mentoring relationship, in contrast, is facilitated through a programme when a young person is paired with a more experienced (usually older) unknown individual with the understanding that the mentor will provide guidance and support that will promote the mentee’s development and over time an emotional connection between the pair will develop (DuBois & Karcher, 2005). Formal mentoring relationships now exist in a variety of forms (DuBois & Karcher, 2005; Karcher, Kuperminc, Portwood, Sipe, & Taylor, 2006; MENTOR, 2009; Weinberger, 2005). *Intergenerational* mentors are exclusively over 55 years, older peers act as mentors in *cross-age peer mentoring*, one or more adults team up with a group of students in *group mentoring*, and *e-mentoring* exists for those who experience barriers to meeting in person (Karcher et al., 2006).

Mentoring relationships occupy a unique position in a young person’s life because they share qualities typical of the parent-child relationship (Beam, Chen & Greenberger, 2002; Keller, 2005; Scales & Gibbons, 1996) but also have qualities in common with relationships with peers (Beam et al., 2002; Keller, 2005). This “hybrid relationship” (Keller, 2005, p. 83) gives mentors the advantage of shifting between these positions depending on the needs of the mentee at any given time. Indeed, mentees have reported that there are certain issues they feel comfortable disclosing to their mentors that they would not divulge to their parents (Beam et
While a similar degree of disclosure may occur in peer-peer relationships, mentees may feel that the advice provided to them by their mentors is more reliable due to their greater experience (Beam et al., 2002; de Anda, 2001). Liang, et al. (2008) proposed that mentor support can assuage some of the pressure adolescents experience in trying to balance their need to become more autonomous whilst still requiring the nurturance found in adult connections.

THE EFFECTIVENESS OF MENTORING PROGRAMMES

The longstanding and pervasive endorsement of mentoring programmes suggests a strong evidence-base for its effectiveness, and the popularity for mentoring programmes is not entirely unwarranted. Positive results have been generated from mentoring relationships across several domains of human functioning (DuBois, Holloway, Valentine & Cooper, 2002; DuBois et al., 2011; Rhodes & DuBois, 2008; Rhodes & Lowe, 2009; Sipe, 2002; Tolan, Henry, Schoeny & Bass, 2008). Agency-facilitated mentoring has been associated with improvements in interpersonal relations (DuBois et al., 2011; Eby, Allen, Evans, Ng & DuBois, 2008; Renick Thomson & Zand, 2010; Rhodes, Grossman & Resch, 2000), school attitudes (LoScuito, Rajala, Townsend, & Taylor, 1996), school attendance (DuBois et al., 2011; Eby et al., 2008; LoScuito, et al., 1996; Rhodes et al., 2000) and performance (DuBois et al., 2011; Eby et al., 2008; Thompson & Kelly-Vance, 2001; Tolan et al., 2008), scholastic competence (Rhodes et al., 2000), school connectedness (Converse & Lignugaris/Kraft, 2009) positive attitudes (DuBois et al., 2011; LoScuito et al., 1996), mental health (Keating, Tomishima, Foster & Alessandri, 2002), problem behaviours (Converse & Lignugaris/Kraft, 2009; DuBois et al., 2002; DuBois et al., 2011; Keating et al., 2002; Tolan et al., 2008) and career-related outcomes (Dubois et al., 2002).

The Big Brothers Big Sisters organisation, for example, has had repeated success in demonstrating the positive effects of their programmes (see De Wit et al., 2007; Maldonado et al., 2008; and Thompson & Kelly-Vance, 2001). These are standalone mentoring programmes that generally target at-risk youth aged
approximately 6-16 years from single-parent households. Rather than targeting a particular problem, the aim is to provide the young person with a supportive and positive role model (Grossman & Tierney, 1998). All mentors undergo screening and obtain training, and the pairs meet several times a month and must remain committed to the relationship for at least one year.

An oft-cited experimental study conducted by Grossman and Tierney (1998) has served as an exemplar in the field as it demonstrated the potential of mentoring programmes to have pervasive effects on youth outcomes. In comparison to a control group (n = 472), youth involved in eight different Big Brother Big Sister agencies (n = 487) were less likely to begin using drugs and alcohol, use aggressive behaviour towards others or lie to their parents. They also perceived greater emotional support from peers; had higher trust in their parents, better school attendance, and academic competence; and slightly better achievement outcomes 18 months after programme commencement.

A randomised controlled trial (n = 562) evaluation of the Across Ages programme demonstrated that successful results can also arise from mentoring when it is part of a multi-component strategy (LoSciuto et al., 1996). Grade 6 youth from high-risk Philadelphia neighbourhoods who were paired with older adult mentors, in addition to participating in community service and life skills training, had better attitudes towards school, their future, and older individuals than youth who only participated in the community and life skills components and a control group that received no intervention. They also scored higher on well-being measures, knowledge of older people and community services, and had better reactions to potential drug use encounters than the control group. Moreover, they had lower instances of substance use than the group that only participated in the other components (LoSciuto et al., 1996).

Nevertheless, several reviews have pointed out that the enthusiasm for mentoring reflected in the proliferation of programmes worldwide (DuBois et al., 2011; Rhodes, 2005, 2008; Rhodes & DuBois, 2008; Rhodes & Lowe, 2008, 2009) and the provision of tremendous amounts of funding (Baker & Maguire, 2005; Rhodes...
& DuBois, 2008; Walker, 2005) does not match the magnitude of the effects demonstrated in the research that is available.

Although supporting the benefits of mentoring for problem behaviours, educational and career-related outcomes and to a lesser degree those related to psychological/emotional functioning, DuBois et al.’s (2002) well-known meta-analysis revealed only small effect sizes of .14 across formal one to one mentoring programmes targeting youth under 19 years of age. Furthermore, DuBois and his colleagues recently conducted a meta-analysis of programmes operating from 1999-2010 and found a similar effect size of .19 (after adjusting for publication bias). Similarly, Eby et al.’s (2008) meta-analysis on the effectiveness of academic, workplace, and youth mentoring across behavioural, attitudinal, motivational, health, and career outcomes also yielded small effect sizes ranging from .03-.14 and the youth mentoring effects were reflected in the lower end of this range. Slightly more promising, Tolan et al.’s (2008) meta-analysis of 39 mentoring programmes targeting delinquent youth or those at-risk for delinquency revealed positive effects for all of the outcomes of interest (i.e. delinquency, drug use, aggression and academic achievement) yet these effects were all modest in size.

It is clear that mentoring is not a cure-all solution. For example, although Grossman and Tierney’s (1998) impact evaluation of the BBBS programme revealed it was an effective programme overall, null effects were obtained for self-concept measures, participation in social/cultural activities, as well as some delinquency and relationship measures. Even when positive effects are obtained they are not necessarily sustained (Rhodes, 2008; Rhodes & DuBois, 2008). Of more concern, however, are the studies that show no improvement on any outcome (see Royse, 1998) and worse yet is the potential for some programmes to produce harmful effects. To illustrate, an experimental study of a school-based BBBS cross-age peer mentoring programme found that when academically connected mentees were matched with mentors who held negative attitudes towards youth in their communities, these mentees were more likely to be negative influences in the classroom than their academically connected control counterparts nine months into the school year (Karcher, Davidson, Rhodes & Herrera, 2010).
That substantial variation exists in programme effects suggests that a small effect size is not an accurate reflection of all mentoring relationships. The combined assessment of relationships that vary in calibre conceals the larger effects that are produced from higher quality matches (Rhodes, 2008; Rhodes & DuBois, 2008; Rhodes & Lowe, 2009; Schwartz, Rhodes, Chan & Herrera, 2010). The same could be noted for the combined assessment of multi-site programmes – cross site variation may exist but without analyses that unveil these differences it is unknown what potential exists for some programme characteristics to produce large effects.

Rhodes & DuBois (2008) have identified a worrying trend that in attempting to satisfy the demand for mentoring programmes, organisations are expanding at a pace that necessitates a reduction in programme intensity and compromises the attention to detail that is needed to be successful. New initiatives are becoming more diverse as they drift further away from traditional models (Rhodes & Lowe, 2008). To prevent wasted resources and the potential for iatrogenic effects, standards for mentoring practice need to be clearly communicated. However, without an understanding of what causes the heterogeneity in the effects obtained, little direction can be provided towards programme development efforts. Ergo, an investigation of the characteristics that distinguish flourishing relationships and programmes from mediocre or injurious ones is paramount and this entails a closer look at programme processes and moderating variables.

**MENTORING PROCESSES AND CONDITIONS PROMOTING GROWTH**

**Mentor approaches and pathways to success**

According to Rhodes and her colleagues, when occurring under the right conditions, discussions and participation in shared activities with a mentor can theoretically cultivate positive growth through mutually reinforcing socio-emotional, cognitive and identity development pathways (DuBois et al., 2011; Rhodes & DuBois, 2008; Rhodes 2005; Rhodes & Lowe, 2009; Rhodes, Spencer,
Emotional well-being can be boosted simply by providing companionship; a mentor who engages in enjoyable activities with their mentee can provide relief from the young person’s daily stressors (Evans & Ave, 2000; Rhodes et al., 2006; Spencer, 2006, 2007a; Spencer & Liang, 2009). Joint involvement in fun activities can also serve as a starting point for developing a strong emotional connection (Converse & Lignugaris/Kraft, 2009; Darling, 2005; Evans & Ave, 2000; Grossman & Bulle, 2006; Liang et al., 2008; Spencer, 2007a, Spencer & Liang, 2009) and a close bond is the crux of a successful mentoring relationship (Converse & Lignugaris/Kraft, 2002; DuBois & Neville, 1997; Parra, DuBois, Neville & Pugh-Lilly & Povinelli, 2002; Rhodes, 2005, Rhodes et al., 2006; Rhodes & DuBois, 2008; Rhodes & Lowe, 2009; Spencer & Liang, 2009). In fact, the results of a path model testing the influence of a variety of youth and mentor characteristics and programme experiences on relationship length and perceived mentee benefits supported a model in which relationship closeness had the only direct effect on these outcomes. Other variables had more distal associations through the mediating role of relationship closeness (Parra et al., 2002).

It is unlikely that this critical sense of intimacy will develop without first satisfying the prerequisite conditions of trust (Rhodes, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009; Sipe, 2002), empathy (Rhodes, 2005; Rhodes & Lowe, 2009; Spencer, 2006), mutuality (Rhodes, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009) and authenticity (Spencer, 2006; Spencer & Liang, 2009). Trust is facilitated when mentors openly share their feelings, communicate their genuine opinions (Spencer, 2006), and create an atmosphere of encouragement and empathy (Spencer, 2007a). This provides a safe arena for youth to more openly express their emotions (Spencer, 2007a). As noted by mentees, a primary function mentors provide is the opportunity to confide in a non-judgmental and trustworthy adult (de Anda, 2001; Spencer, 2006, 2007a; Spencer & Liang, 2009). Acting as a “sounding board” (Rhodes & Lowe, 2009, p.172), mentors can provide new perspectives and useful advice (Rhodes, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009, Spencer, 2007a).
which can facilitate the regulation of emotions and the development of communication skills (Evans & Ave, 2000; Rhodes, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009; Spencer, 2006, 2007a). A qualitative study of natural mentor relationships in Scotland revealed that both mentors and mentees identified a broker role as an important function of mentors which often entailed allowing mentees to test out their ideas or disclose sensitive information to their mentor before relaying it to the actual targeted individual (Philip & Hendry, 2000).

A young person’s competencies can be markedly improved when learning is approached as a mentor-mentee collaboration (Liang et al., 2008; Spencer, 2006; Spencer & Liang, 2009). Cooperative tasks and assistance from others allows individuals to stretch their capabilities and accomplish things that would not be attainable on their own (Rhodes, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009). The zone of proximal development (ZPD) “is the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers.” (Vygotsky, 1978, p. 86). Vygotsky (1978) felt that working within the ZPD was the key to the development of new cognitive processes because with guidance or collaboration these capabilities become internalised and subsequently reflected in the individual’s actual developmental level.

Care must be taken not to be overly-directive, however. A mentor who prescribes activities, determines the young person’s goals, and imposes his or her own values is likely to alienate his or her mentee (Dunphy et al., 2008; Morrow & Styles, 1995). Successful mentoring relationships tend to be those that are youth-centred and autonomy supportive—showing the young person respect and allowing their input into decisions (Morrow & Styles, 1995; Sipe, 2002). Liang et al. (2008) found that this was especially valuable for youth in high school or college as compared with those in middle school.

Participation in new activities introduced by a mentor can also play an important role in identity development. It is often in trying new activities that new
interests bud and when one’s strengths and shortcomings become evident. (Rhodes et al., 2006; Rhodes & Lowe, 2009). The young person may also begin to identify with the new peer groups that these settings provide.

Consistent positive experiences with an adult mentor have the potential to alter any pre-existing pessimistic views of adults that may have been shaped through previous unpleasant encounters (Rhodes, 2005, Rhodes et al., 2006; Rhodes & Dubois, 2008; Rhodes & Lowe, 2009) making the young person more receptive to intergenerational relationships. Research has demonstrated that mentoring relationships can influence disclosure to and friendships with other adults (Renick Thomson & Zand, 2010). This has important implications because interactions with different adults yield a comparison base for what one may aspire to be (and be like) in the future (Jarrett, Sullivan, & Watkins, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009; Spencer & Liang, 2009).

Programmatic Influences

It is clear that a mentor’s approach has important implications for the development of a high quality relationship but the onus must not be completely placed on the mentor; programmatic conditions and policies play a big part in facilitating adult-youth connections (Grossman & Bulle, 2006). Grossman & Tierney (1998), for instance, attributed the success of the BBBS programme to the high infrastructure of the programme.

Among the best practices outlined for programme policies is the communication of expectations for regular and consistent interactions and the requirement of a one year commitment (MENTOR, 2009). More frequent contact between mentors and mentees has been associated with increased interpersonal closeness and greater mentee benefits (DuBois & Neville, 1997). Moreover, relationships that terminate early can cause adverse effects for mentees (Grossman and Rhodes, 2002). Unrealistic expectations and a lack of skills on the part of the mentor are also cited as sources of early relationship termination (Spencer, 2007b; Morrow & Styles, 1995) therefore other agency best practice standards include screening mentors, providing pre-match and ongoing training as well as ongoing
supervision of the relationship (DuBois et al., 2011; Karcher et al., MENTOR, Sipe, 2002; Dunphy et al., 2008). Youth may initially approach relationships with adults with discomfort, but activities that promote interactions allow young people to see adults in a new light and this can cause a shift to a more positive perception (Jarrett et al., 2005). This may be why agencies that organise opportunities for structured (DuBois et al., 2002; Parra et al., 2002) and group (Chan & Ho, 2008) activities also tend to have better results.

The criteria for matching the mentor/mentee pair are also relevant to programme success. DuBois et al., (2011) discovered that programmes matching pairs based on similar interests were more effective than others and under natural circumstances, young people tend to select mentors who are similar to them (Liang & Grossman, 2007). Many agencies prefer to match mentees with a mentor of the same gender (all BBBS programmes for example) and racial or ethnic background (Liang & Grossman, 2007; Rhodes, Reddy, Grossman & Maxine Lee, 2002). Theoretically, cultural mistrust and beliefs that one will be negatively stereotyped can reduce a mentee’s receptivity to a mentor of a different race (Darling, Bogat, Cavell, Murphy & Sanchez, 2006) and positive role models of the same ethnicity are thought to be especially important to the identity development of ethnic minorities (Evans & Ave, 2000; Sipe, 2002). This may be particularly pertinent for young Maori in New Zealand (Dunphy et al., 2008) because some researchers believe that the effects of colonial oppression continue to engender disempowering self-beliefs (Hollis, et al., 2011; Puketapu-Andrews, 1997).

Unfortunately, the societal barriers experienced by ethnic minorities reduce the capacity for these individuals to become mentors (Dunphy et al., 2008) leading many to opt for cross-ethnic matching as opposed to providing no mentor at all (Rhodes et al., 2002). While differences in cultural backgrounds can be challenging for a mentor (Morrow & Styles, 1995; Spencer, 2007b) and at times lead to the failure of the relationship (Spencer, 2007b), cultural competency training and an empathetic approach can often remedy these tensions. What appears to be more important is that the pair has similar interests and that their preferences are taken into consideration (Sipe, 2002; Spencer, 2007b;) and actually, an investigation of the
differential effects of cross versus same-race matching in American BBBS programmes revealed no consistent negative or positive results across outcomes for either pair configuration (Rhodes et al., 2002). Furthermore, several studies have found no evidence that ethnically congruent pairs fare better than cross-ethnic matches (DuBois et al., 2002; Sipe, 2002; Parra et al., 2002).

**Mentee Characteristics**

Taking into account the developmental stage of adolescent mentees, it may not be appropriate to expect mentees to take responsibility of the relationship and initiate contact in the initial stages of the relationship (Morrow & Styles, 1995; Sipe, 2002; Spencer, 2007b). Mentors need to be patient and persevere with maintaining the relationship when mentees seem disinterested (Rhodes et al., 2006) but without eventual mentee engagement and commitment, growth will be inhibited (Dunphy et al., 2008; Spencer, 2007b). The mentor and programmatic factors delineated above should influence the degree to which a mentee engages with the relationship but the mentee’s unique characteristics, previous relationship experiences, and social location with regards to socio-economic conditions and associated risk factors may also contribute to the heterogeneity in relationship success (Darling, et al., 2006; Schwartz et al., 2010). Below, I explore three of these factors: gender, ethnicity and environmental risk status.

**Gender differences**

Female relationships tend to be more intimate in nature leading some researchers to feel that a psychosocial approach to mentoring that focuses on developing an emotionally close relationship is more appropriate for girls whereas boys may prefer an instrumental style that is activity-focused (Darling et al, 2006; Liang & Grossman, 2007).

In their qualitative study of 82 BBBS relationships, Morrow and Styles (1995) pointed out that males experienced greater enjoyment when the partnerships were characterised by shared activities whereas females were more likely to emphasise talking and personal disclosure. In spite of this, an in-depth exploration of
processes characterising female mentor-mentee relationships conducted by Spencer and Liang (2009) found that although females emphasised the importance of emotional intimacy in their relationships, an instrumental role was also evident and valued. In a similar study of male mentoring relationships, Spencer (2007a) demonstrated that a primary theme was one of emotional support and connectedness.

Mixed results have been obtained in other studies investigating gender differences in mentoring. Grossman and Tierney’s (1998) BBBS sample was used once again to investigate gender differences in relationship characteristics and length. Rhodes, Lowe, Litchfield, and Walsh-Samp (2008) found that female partnerships lasted longer; females reported more helpful behaviour from their mentors in the longer lasting relationships; and were more satisfied with longer relationships than males—who, in contrast, were more satisfied when relationships were of short or medium duration. In contrast, Morrow and Styles’ (1995) study revealed that males were more commonplace in the longer lasting and higher quality partnerships but noted that referrals to the programme differed for each gender: males were selected because they lacked a positive male role model while females were generally selected because of emotional or behavioural problems. In line with this, DuBois et al.’s (2011) more recent meta-analysis revealed that programmes serving a higher proportion of males were more effective. However, the early meta-analysis conducted by DuBois et al. (2002) did not find mentee gender to be a moderating factor of programme effects nor did Parra et al. (2002) and Keating et al. (2002) find any gender differences in relationship characteristics or outcomes.

**Ethnic differences**

Above we saw that discrepancies in the ethnic background of the pair can create additional complexity in the relationship which ultimately can impact on programme success. As such, research on ethnic differences in mentoring relationships has focused largely on the “matching” issue (see Darling et al., 2006 and Liang & Grossman, 2007). Cultural values also have implications for the
agency-determined structure of mentoring relationships. Youth from ethnic backgrounds that prioritise family and community interdependence and group success (collectivism) over self-reliance and individual achievement (individualism) may not be as receptive to a one to one mentoring structure emphasising the young person’s autonomy and his/her individual goals (Liang & Grossman, 2007). Evans & Ave (2000) also warned about the possibility that, despite good intentions, the skills and values promoted by a mentor may actually be maladaptive for the young person’s familial or cultural context.

The level of family involvement can moderate the effectiveness of mentoring programmes (DuBois et al., 2002) and this may be especially so for individuals from collectivist cultures (Darling et al., 2006; Liang & Grossman, 2007). Alternative suggestions to the traditional model include: mentors working with the whole family; using group mentoring configurations; using multiple mentors (Darling et al., 2006; Liang & Grossman, 2007); and the development of programmes that support the creation of mentoring relationships within the young person’s existing social network (Evans & Ave, 2000; Evans, Jory, & Dawson, 2005; Liang & Grossman, 2007).

Like gender, results have been mixed when focusing on the differential effects of a programme across ethnic groups. Keating et al.’s (2002) evaluation of a US programme targeting youth at-risk for delinquency or mental illness showed that African Americans were affected differently than non-African Americans (mostly Caucasian and Latino). While mothers of non-African Americans reported decreases in maladaptive behaviours after the programme, levels remained approximately the same for African Americans (one should note that the pre-programme scores between these groups were not equivalent). Grossman & Tierney (1998)’s impact evaluation of BBBS programmes also demonstrated differences in the strength of effects between ethnic minorities (primarily African Americans and Hispanics) and Caucasian mentees. For example, the post-intervention decrease in drug use was largest for ethnic minorities.
Socio-economic Location and Risk Factors

As outlined previously, New Zealanders living in areas of high deprivation experience worse outcomes across educational, vocational, and health domains (Ministry of Social Development, 2010; Adolescent Health Research Group, 2008), which is likely linked to reduced access and exposure to important determinants of health and well-being (Te Röpū Rangahau Hauora a Eru Pōmare, 2002). One theorised function of mentoring programmes is, therefore, to compensate for the lack of opportunities (educational, recreational, support, etc.) available to young people living in areas characterised by high socio-environmental risk (Evans & Ave, 2000). In fact, the great majority of mentoring programmes in the US target at-risk youth (DuBois et al., 2011; Jekielek, Moore, Hair, & Scarupa, 2002). For example, the Big Brothers Big Sisters programmes stipulate that potential mentees come from single-parent households, the Brothers Project (Royse, 1988) and Project R.E.S.C.U.E. (de Anda, 2001) accommodated youth living in neighbourhoods characterised by poverty and crime as does the Across Ages programme (LoScuito et al., 1996).

Socio-environmental risk factors can compound the responsibilities of an adult mentor. Mentees who are situated in contexts abound with violence, poverty, substance use, and familial discord mean that mentors may need to take on stressful “crises-management” roles (Ginwright, 2005, p. 103). Coupled with this is the pattern of youth living in high deprivation areas taking on additional responsibilities at home (Adolescent Health Research Group, 2008; Ginwright, 2005), often to compensate for the longer hours parents (or single parents) work (Ginwright, 2005). This may require mentors to re-think the commonly recommended approach of affording mentees more responsibility and sharing decision-making. Ginwright (2005) explained how some adolescent girls he interviewed expressed a desire to have a reprieve from making decisions and having responsibilities.

Because all participants in many programmes are considered at-risk, studies often do not include comparisons across risk status and/or socioeconomic status
(SES) groups. DuBois et al (2002), however, did assess whether these factors moderated programme effects in their meta-analysis. Programme effects were contingent on the at-risk statuses of the included samples with stronger effects for programmes serving participants with high environmental risk, and those for participants with both individual (e.g. behavioural problems) and environmental risk, but not programmes targeting individual risk alone. A nearly significant effect was also obtained for SES status indicative of more favourable effects for low SES groups. In their more recent meta-analysis DuBois et al. (2011) examined the interaction of individual and environmental risk on programme outcomes. They discovered that programmes were more effective when environmental risk was high but individual risk was low and conversely when individual risk was high but environmental risk was low, but not when both were high.

In an effort to assess whether the disadvantages apparent in urban schools (i.e. stretched resources, less experienced teachers, students from poverty-stricken communities) would create barriers to the delivery of successful mentoring programmes, Dappen and Iserhagen (2006) assessed differences in the degree of positive behaviour change exhibited by urban and nonurban students following involvement in a school-based, adult-youth mentoring programme based in a midwestern US state. Again, mixed results were obtained: students and mentors from urban schools rated more positive behaviour than those from nonurban schools, while no differences were detected in parent and teacher ratings.

To sum the above sections, one may reasonably expect differences in the nature and quality of the relationship based on the characteristics participants (and mentors) bring to the table. However, there is no conclusive or even consistent pattern of differential programme effectiveness based on gender, ethnicity, or risk status (environmental or other). Clearly, more research is needed to uncover what occurs behind the scenes of these dyadic interactions when different subgroups of individuals are involved.
MENTORING IN THE NEW ZEALAND CONTEXT

Learning through mentorship is an integral part of the indigenous culture in Aotearoa/New Zealand. Older siblings or cousins (tuakana) are expected to take their younger sibling/cousin of same gender (the teina) under their wing and guide their learning and development (Tangaere, 1997). Whānau elders are also critical in shaping the identities of young Māori, assuming responsibility for their language and cultural education (Dunphy et al., 2008; McCarthy, 1997; Tangaere, 1997). Likewise, in traditional times, practical skills were often refined through an apprenticeship style of education (Tangaere, 1997).

Ironically, the framework for formal youth mentoring programmes was not derived from local practices but rather mimicked US schemes (Evans & Ave, 2000; Evans, Jory, & Dawson, 2005). The first programme was established in the early 1990’s (Farrugia et al., 2010) and, paralleling the mentoring movement in America, programmes have been on the rise ever since (Dunphy et al., 2008; Evans & Ave, 2000; Evans, Jory, & Dawson, 2005; Farrugia et al., 2010). Unfortunately, evaluation research on mentoring in New Zealand has not kept pace with the expansion of its practice (Farrugia et al., 2010) – not unlike the trend in the US. In their article (published in 2000) Evans and Ave claimed “there are as yet no published, methodologically-adequate, outcome studies of mentoring projects in New Zealand” (p. 44).

This is of concern because although there is evidence to support its effectiveness overseas, US models of mentoring may not provide the best cultural fit for New Zealand society (Evans & Ave, 2000; Evans, Jory, & Dawson, 2005). Māori, Pacific, and Asian peoples account for approximately 40% of young people aged 10-19 (Statistics NZ, 2006). Each of these groups is considered a collectivist culture and some concerns about the applicability of the one to one mentoring model to these cultures have already been delineated above. Mentoring has been recommended as a strategy to improve the outcomes of both Māori (Clark et al., 2008) and Pacific youth but the Pacific Youth Development Strategy stressed that mentoring strategies should involve community elders and attend to the needs of
Pasifika youth (Ministry of Social Development, 2005). Dunphy et al. (2008) suggest that non-Māori mentors can be involved in reconnecting young Māori to their cultural identity, but this requires attention to New Zealand’s historical influences and an awareness of one’s own cultural values and position of power. Others have cautioned that organisations should do more to incorporate indigenous concepts in their programmes (Evans & Ave, 2000; Evans, Jory & Dawson, 2005).

Single-parent families in Māori culture, for example, do not necessarily have the same connotations of risk as they do in other western societies (Evans & Ave, 2000); the whānau network shares the responsibility of raising young people in Māori communities (Evans & Ave, 2000; McCarthy, 1997). This also means the extended family may want to play a role in selecting which mentors are paired with their younger whānau members (Evans, Jory, & Dawson, 2005). According to Evans, Jory, & Dawson (2005), the U.S. notion of mentoring which insinuates the acceptance of support from a stranger, who presumably holds more social power, may also be at odds with some core New Zealand values. They posit that New Zealanders pride themselves on being able to rely on friends (mateship) but also treasure their ability to be self-sufficient. Individuals who display an air of superiority are look upon with derision as there is a strong belief that all are entitled to a “fair go” – or equal opportunities, and this may affect the receptiveness of some young people to particular mentors. Bidirectional teaching and learning roles are also a feature of the tuakana/teina relationship. It is not uncommon for the teina to take on an instructional role at times (Tangaere, 1997). In consequence, mentoring practices in New Zealand should account for these cultural nuances.

The near absence of published New Zealand-based mentoring programme evaluations means little is known about which practices are most effective in this context. Fortunately, mentoring research has recently attracted government attention. As part of their “What Works in Youth Development” series, the Ministry of Youth Development commissioned a systematic review in 2009 to
assess the effectiveness of youth mentoring in the New Zealand (Ministry of Youth Development, n.d.).

Findings from this review are encouraging with 88% of the included programmes demonstrating effectiveness at some level (Farrugia et al., 2010). A look at some of the moderating factors indicated that programmes were more successful in influencing the psychosocial and interpersonal domains as opposed to educational, behavioural, vocational or cultural outcomes. Programmes that were more structured, and/or had been operating for a longer duration, and/or had a history of evaluation fared better than others as did those that were part of a multi-component strategy and those that incorporated principles of best practice. In terms of participant demographics, programmes targeting participants from mid-range SES backgrounds tended to be less effective than those at opposite ends of the SES continuum but age, gender and individual risk status specific programmes did not differ across subgroups (Farrugia et al., 2010).

Despite the notion that group mentoring may be more fitting for Māori, Pacific and Asian cultures, 72% of the programmes in this review facilitated one to one relationships. Moreover, although 14 programmes served Māori participants only half of these could be said to have attended to any cultural factors. Only six programmes included Pacific participants and of these one was rated as having low cultural appropriateness and another as having ignored Pacific culture; however the majority accommodated for the general New Zealand culture (Farrugia et al., 2010). This finding appears to contradict statements made by The Youth Mentoring Trust that mentoring programmes have an appreciation of the cultural diversity within the younger sector and there is an increasing focus on programmes that are appropriate for Māori and Pacific peoples (Dunphy et al., 2008). It is also disappointing to note that although Asian young people constitute an equivalent proportion of the youth population as Pacific youth, there is virtually no mention of this group in mentoring research or policy documents.

Surprisingly, programmes that used group mentoring tended not to be as effective as the one to one or mixed structures and programmes that were more
culturally appropriate for Māori were also less effective. There was no difference between those programmes that attended more closely to Pacific culture and those that did not. Furthermore, family involvement did not contribute to variation in programme effectiveness. In the case of Māori culture, the authors of the youth mentoring systematic review speculate that these programmes may not have attended carefully to the mentoring best practice principles, but they emphasise that these are not incongruent with the Māori cultural practices and values embedded in these programmes (Farruggia et al., 2010).

The review also provided a clearer picture of New Zealand efforts in this particular research area, and this picture rather disconcerting. Twenty-six studies of the 74 that were considered ended up being reviewed. Of these only eight were evaluations of currently active programmes. Communications with the Youth Mentoring Trust indicated that 23 programmes were in operation at the time of the review; thus only 35% of these had conducted an outcome evaluation and only 13% had measured outcomes 6 months or more after the intervention. (Farrugia et al., 2010). Additionally, the quality of the research within the 26 studies was extremely variable which necessitated the use of fairly lax criteria for determining the type of research methodologies to be included; thus the results are tenuous at best. To illustrate, qualitative studies were accepted if they indicated some kind of change on an outcome and some studies with only post-test measures were included (Farrugia et al., 2010). According to their systematic review process, Project K’s outcome evaluation is the only randomised controlled trial of a mentoring programme in New Zealand.
COMMON YOUTH PROGRAMMING THEMES

There is substantial thematic overlap across these three genres of youth development programmes. The commonalities between adventure, service-learning, and mentoring programmes are found in aspects of their rationales, goals, processes, results, ideal programme conditions, and areas for future research. Each attempts to address some of the social deficits brought about by the major societal changes that have occurred since at least the middle of the twentieth century. All three are also particularly strong in effecting positive changes in self-concept and social development. Perchance this stems from their dual emphasis on nurturing individual strengths within a cooperative framework. Perhaps this is why each has gained so much popular support in the past few decades.

At a more detailed level, these mutual strengths may emerge from the similar supportive approach taken by the facilitators and mentors in the programmes. Participant mastery is cultivated through the use of scaffolding: adults support the autonomy of the individual participants, giving them voice and ownership over their own learning journeys, while setting the stage for success by providing core guidelines and remaining accessible for informational or instrumental support when it is required. Moreover, participant growth strategies are embedded in an experiential framework where authentic activities, observation, modelling, reflection, and feedback are highlighted. The exploration of new interests and activities is also a common aspect across the three classes of programmes but each promotes this aspect of identity development via different avenues.

The results from effectiveness evaluations of adventure, service-learning, and mentoring programmes are tilted towards the positive but the literature across the board is characterised by small to moderate positive effects and often results are mixed. The variation in programme effectiveness appears to reside in the differences across programmatic conditions and in individual participants. The reviews of all three programme types point to the importance of adequate dosage
and attention to individual differences, although no consistent conclusions can be made for differential effectiveness across gender, ethnic identities, or socio-economic location and external risk factors.

This heterogeneity in programme effects has provoked calls from researchers in each programme field for more investigations into the processes and moderating factors which influence success in adventure, service-learning, and mentoring programmes for youth. It is clear that the research knowledge base for each of these youth development areas would be strengthened by the inclusion of more rigorous research studies, including longitudinal research and mixed method designs.

**DISTINCTIVE FEATURES**

Although one may find consistent emphasis of the aforementioned features across the three programme types, each has unique strengths and limitations which differentiate them from each other. Adventure programmes capitalise on the growth-promoting dissonance found in an unfamiliar social and physical environment and the risk and uncertainty of the challenges that these conditions generate. Nevertheless, this field has been criticised for the very real physical risks inherent in the programme activities and for a lack of attention to emotional safety. Service-learning programmes have a distinctive focus on civic responsibility and sense of community, which consequently, produces direct benefits to a society. On the other hand, success for these programmes is contingent on positive collaborations with community members who are not directly tied to the programme and on attentiveness to youth agency. Otherwise, the absence of these elements can potentially drive a greater wedge between the young people and adults of a community. Mentoring programmes provide nurturance and emotional support at an intensity that is not found in the other two kinds of programmes, possible in part because of the undivided attention an individual participant obtains from a one-to-one relationship. Individuals with the best intentions, however, may realise that a one year commitment is too much and terminating a
relationship early can reinforce a young person’s already negative self-concept or negative view of adults as dependable support figures.

**THE STATE OF AFFAIRS IN NEW ZEALAND**

The meagre supply of New Zealand specific research in the three fields tends to mirror the positive trends found in overseas investigations, but as the context-specific evidence base is so thin a caution is warranted to those looking to use the findings to inform practice. In addition to the previously mentioned areas for general future research growth, even effectiveness evaluations (which are much more ubiquitous in the general evaluation literature) are sorely lacking across the three programme areas in New Zealand. Adventure programming is the least concerning in this regard because results from commercial programmes, such as OBNZ, and school-based programmes have been published in peer-reviewed outlets; however, the studies that I was able to access usually employed small samples and lacked comparison groups of any kind. Moreover, Farruggia et al. (2010) underscored the poor quality and the limited number of youth mentoring programme evaluations in their systematic review of the New Zealand-based mentoring literature and I could not source even one study of a service-learning programme for young people in New Zealand.

The small headway made in the adventure programming research field may be attributable to the longer existence of these kinds of programmes in New Zealand and this is perhaps because this particular programming philosophy aligns closely with the nation’s reverence for outdoor adventure. In contrast, mentoring and service-learning programmes have not had the same history of use and, as they are currently framed, are American imports.

Nevertheless, there has been noted growth of these programmes in New Zealand, particularly adventure and mentoring programmes and the reality seems to be that many programme developers are uncritically implementing overseas programme models. Yet there is a clear need to carefully consider how these
models fit with the unique multicultural make-up of New Zealand society as well as how they may or may not reflect important bicultural policies.
Chapter Four
CONSTRUCTION OF THE WHITE BOX
PROJECT K’S THEORY OF CHANGE

If you don’t know where you’re going, you might not get there ~ Yogi Berra
There is nothing so practical as good theory ~ Kurt Lewin

THEORY-DRIVEN EVALUATION

As was expressed in Chapter Two, third stage evaluation theories are those that attempt to reconcile the limitations of the first and second stage approaches (Donaldson, 2003, 2007; Shadish et al., 1991). This is largely due to concerted attempts to address both the problem of black box evaluations endemic to the earliest approaches while also redressing the reputation that the utilisation-focused approaches of the second stage lack scientific rigor (Shadish et al., 1991). According to Shadish and his colleagues (1991) the theory-driven approach (and the family of models that exist under this umbrella) falls under the third stage classification. The overarching goal of the theory-driven approach is to use rigorous methods to unearth the mechanisms thought to be driving the positive change that occurs as a result of the social programme under investigation; hence, the “white box” moniker that is sometimes tagged to it (Astbury & Leeuw, 2010, p. 364).

Theory-driven evaluation has gained a long list of advocates over the last 40 years. Even Thomas Cook, a staunch advocate of experimental or quasi-experimental design evaluations, endorses the utility of theory-driven approaches (Cook, 2000). Cook goes so far as to say that the combination of a theory-driven approach within a randomised experimental paradigm is ideal for programme evaluation. Perhaps we are moving towards a new “golden standard” that involves producing convincing findings about causal attribution while simultaneously penetrating the black box.
CHARACTERISTICS OF A THEORY-DRIVEN APPROACH

The rise in popularity of theory-driven approaches has generated numerous adaptations to the foundational ideas but a discussion of such nuances is beyond the scope of this chapter. Interested readers are directed to the works of some of the more prominent theory-driven advocates such as Bickman (1989), Chen (1990) and Rossi (Chen & Rossi, 1980, 1983, 1987); Donaldson (2003, 2007), Rogers (2000, 2007), Weiss (1997a, 1997b, 1998b) and Wholey, 1987). Across the many versions of this approach, the two defining aspects are: 1) making explicit the rationale which details how the programme is thought to produce the desired outcomes and 2) using that rationale to guide the next steps of the evaluation (Rogers, 2007).

The first step of developing the programme theory generally involves interviews or discussions with key programme stakeholders. This reveals their tacit assumptions of how the programme operates. Also, a collaborative and interactive approach, as was suggested in Chapter Two, usually increases the likelihood that the evaluation findings will be used (Fetterman, 2003, Fetterman & Wandersman, 2004; Patton, 1997; The Health Communication Unit, 2001; W.K. Kellogg, 2004). However, a more comprehensive approach uses multiple sources (e.g. observations of the programme in action and reviews of key programme documents, prior evaluations of the programme or similar initiatives, and the relevant academic literature) (Donaldson, 2003, 2007; McLaughlin & Jordan, 1999; Scheirer, 1987; Weiss, 1997b). These perspectives are then compiled into one overarching coherent theory.

One challenge of using multiple sources is that competing pictures of the programme theory may result, but understanding the degree of consensus across the organisation is valuable information in and of itself. This may prompt further dialogue or actions to reconcile diverging views. Alternatively, if consensus cannot be reached then the competing theories can be tested (Weiss, 2000).

The final theory or theories are then usually depicted visually as a tabular or graphic programme logic model. Donaldson (2007) recommends that this representation be a more parsimonious version of the more detailed theoretical
explanation that accompanies it. Depending on stakeholder preferences, evaluator skills, the length of time, and allocated resources, the theory is then tested. Methodologies could include detailed qualitative inquiries, causal modeling, basic two-step modeling, stage-state analysis, or developing substantive models (see Lipsey and Pollard, 1989 for details on the latter three). One should note that most programmes are far too complex to allow a test of the comprehensive theory; thus, decisions about which links to investigate must be made (see Weiss, 2000 for recommendations on how to approach these decisions).

**BENEFITS OF A THEORY-DRIVEN APPROACH**

Numerous benefits result from this approach. The most immediate of these is *evaluative direction* (Donaldson, 2003; Lipsey & Pollard, 1989; Weiss, 1997a). The program theory development process should point to the key theoretical variables that generate short-term outcomes and important moderating variables. The process can also indicate when one can expect to observe certain outcomes and, at times, it can reveal unintended side effects that may have gone unnoticed (Donaldson, 2003; Lipsey & Pollard, 1989). Once the bigger picture and the critical variables have been outlined, the evaluator should have a better idea of what methodological design will be the most useful for subsequent evaluations.

Understanding who to focus on, which variables to focus on, and when to focus on them will also *improve the sensitivity of the evaluation design* (Donaldson, 2003, 2007). Often outcome evaluations are developed around stated objectives which are unrealistic or unlikely to occur until the distant future. This possibility is reduced if theory-driven evaluation is used. It also reduces the chances of finding null effects due to insensitive evaluation designs that fail to account for other extraneous sources of variance (Cook, 2000; Donaldson, 2007).

The theory-generating process can also *promote organisational learning and development* (Donaldson, 2003, 2007), especially in circumstances where multiple stakeholder groups have been included. The process can reveal differing conceptualisations of the programme within the organisation and it can clarify if the programme activities are being delivered in the same manner across the board.
Many describe it as a consensus-building exercise during which a common understanding emerges (Kaplan & Garrett, 2005; The Health Communication Unit, 2001; W.K. Kellogg Foundation, 2004). Plus stakeholders gain valuable programme development information about programme implementation. This information can increase their sense of ownership over the evaluation findings (W.K. Kellogg Foundation, 2004).

Furthermore, the programme logic model output, which is generally a one-page summary, is an ideal communication tool to convey the underlying rationale of the programme to external stakeholders such as potential funders and clients (Donaldson, 2007; Kaplan & Garrett, 2005; McLaughlin & Jordan, 1999; W.K. Kellogg Foundation, 2004).

Finally, many advocates (e.g. Astbury & Leew, 2010; Chen, 1990, 1994; Chen & Rossie, 1980, 1983, 1987; Donaldson, 2003, 2007; Lipsey & Pollard, 1989; Rogers, 2000, 2007; and Weiss, 1997) argue that theory-driven evaluation can contribute to a cumulative knowledge base about similar social programs because often the mechanisms producing positive changes and the moderators that condition the processes are common across a number of approaches (Weiss, 1997a).

**A WHITE BOX EVALUATION OF PROJECT K**

In the current study a randomised controlled trial of the programme was already underway (the focus of Chapter Five) and so the question of overall effectiveness was being addressed. Given that the evaluation was the basis of a doctoral thesis, the timeline was lengthy. Because I conduct evaluation research within an academic institution, resources were plentiful as was access to skilled social scientists. Furthermore, it was necessary that the research make an original contribution to the field and developing advanced statistical and analytical skills was a personal aim. It was especially important to contribute to the New Zealand youth development knowledge base because systematic evaluative research specific to the region was sorely lacking. Towards these ends, a theory-driven approach was a very suitable choice.
However, the critical reason for adopting any evaluative strategy is that it will address the information needs of the primary intended users. In previous work with FYD and Project K (i.e. the evaluation questions voting process, see survey in Appendix F), programme stakeholders indicated that understanding why, how, for whom and under what conditions the programme works were high priority questions. I also hoped that the theory-generating process would increase stakeholder engagement in the evaluation and produce more useful results, as there was evidence that the difficulties encountered while implementing the randomised controlled trial had produced substantial resistance to evaluation across most of the programme regions (as I pointed out in Chapter One).

Additionally, it appeared that the organisation could benefit from a consensual understanding of the programme theory as the programme had been in operation for over a decade, it had been scaled up to a large degree, and there had been some staff turnover. Finally, because theory-driven evaluation emphasises rigorous methods and collaboration with stakeholders, it is an approach that can clearly address the dual aims of the *Thoughtful Evaluation* model.

**THEORETICAL INFLUENCES OF THE CURRENT APPROACH**

My approach drew primarily on the theory-driven models of Weiss (1997a, 1997b, 1998b), and Donaldson (2003, 2007). It also incorporated aspects of Gugiu and Rodriguez-Campos’ (2007) Semi-Structured Interview Protocol (SSIP) for building programme logic models and the direct logic analysis process described by Brouselle and Champagne in 2011. To clarify, I sought to develop a logic model and associated programme theory that would focus on the mechanisms of change and the moderators of programme effects (e.g. implementation aspects that can influence the success of the programme, such as gender or programme length). Weiss (1997a, 1998a) defines this combination as a programme’s *theory of change*. I included the “theory development”, “prioritising questions”, and “answering evaluation questions” steps of Donaldson’s (2003, p. 114, 2007, p. 10) three-step *Program Theory Driven Evaluation Science* approach; however, the evaluation questions were prioritised (using the Evaluation Questions Voting Survey in Appendix F) prior to developing the programme theory to ensure that I obtained
stakeholder input before deciding if a theory-driven approach was appropriate. I used a careful selection of Gugiu and Rodriguez-Campos’ (2007) comprehensiveSSIP. More specifically, I focused on the primary target group, the social context, and the individual-level processes and outcomes. After the finalised programme logic model was produced, a process akin to the direct logic analysis described by Brouselle and Champagne (2011) was conducted. A logic analysis involves comparing stakeholder conceptualisations of a programme against recent and/or foundational academic literature relevant to the programme under investigation and/or to expert opinions.

THE CURRENT APPROACH

Research Aims

The primary research aim of this study was to develop a programme theory and associated logic model grounded in the programme staff and founder experiences. This theory would detail the proposed processes of change that individuals undergo as a result of Project K and could be used to guide future empirical investigations of the theory. Secondary aims were to assess the plausibility of the proposed model against a social science literature base and to build on the theory-driven evaluation literature by producing a detailed example of a theory-building process that meets the standards of scientific rigor but also empowers stakeholders, and is feasible within a real-world context.

METHODOLOGY

THE THEORY-BUILDING PROCESS

Incorporating the above theoretical influences, I developed a staged theory-building process. The purpose of the first stage was to build a consensual theory of the mechanisms presumed to be producing the desired outcomes in Project K. I labelled this the Exploratory Stage of Theory Development as I was exploring thoughts about the programme provided by a few key stakeholder groups. I also drew a distinction between two steps within this stage:
1. The first step involved generating the preliminary theory which was grounded in the perspectives of the programme staff and the founders (who are also National Support Office staff). I began here because these stakeholder groups had the most intimate knowledge about how the programme was conceptualised when it was first launched and how it actually unfolds in the here and now.

2. The second step involved assessing the level of consensus across several different programme sources by comparing the information provided in each. Specifically, the sources included: focus groups with staff associated with Project K (the primary source mentioned above); open-ended responses of Project K youth participants provided in an end of programme survey; key programme documents; and previous research studies investigating Project K.

The second stage was labelled the Validation Stage of Theory Development. It involved evaluating the legitimacy of the claims proposed in stage one with more objective sources. Again, I distinguish between two steps:

1. Assessing the surface legitimacy\(^{10}\) of the programme theory by assessing whether or not the proposed links between the programme processes, influencing factors (i.e. moderators), and participant outcomes seemed plausible according to well-established social science theories. This was done by reviewing the relevant research literature (presented in Chapter Three) and the process was more or less equivalent to the direct logic analysis process described by Brouselle and Champagne (2011).

\(^{10}\) I described this step as assessing “surface” legitimacy because this stage consisted of a qualitative comparison of the proposed theories with well-established theories of social science but no empirical evidence was collected; whereas, the focus of the subsequent step would focus on doing the latter and thus would test the legitimacy in terms of statistical properties.
2. *Testing the hypothesised links* was done in part\(^{11}\) by assessing some of the data collected as part of the randomised controlled trial outcome evaluation (see Chapter Five). Additionally, we (FYD’s Evaluation Coordinator and I) used the theory developed in the first stage to provide direction with identifying other potentially important variables. As a result we sourced additional measures to test a few of the proposed causal sequences in a follow up study. This study is on-going and thus not included in this thesis.

Figure 3 summarises the stages of the theory development and the sources, methodology, and the general analytical strategy associated with each. The focus of this chapter is to outline the process and results of the *Exploratory Stage* and the first step (i.e. assessing the surface legitimacy) of the *Validation Stage*. The results for the *Exploratory Stage* are provided within distinct sections for each different data source. Any new or diverging ideas from the staff-driven preliminary model are noted in the relevant comparative data source results section. A general discussion follows which summarises the overall fit of the preliminary model across all of the data sources; the final Project K logic model is then presented and its congruence with the relevant academic literature is explored.

---

\(^{11}\) Because practical constraints impose limitations on the number of causal links an evaluator can investigate at any one time (Weiss, 2000), one must be selective about the questions pursued and a responsive evaluator will be attentive to which questions and methods are most appropriate within the specific context; hence, the *testing hypothesised links* step cannot realistically be completed within one study and presumably could go on indefinitely as new ideas for theoretical links emerge.
Figure 3. The Programme Theory-Building Process ordered by stages and steps according to Project K data source, method type, and analytic strategy.
EXPLORATORY THEORY DEVELOPMENT: GENERATING THE THEORY

Pilot Process: Workshop with National Support Office Staff

A programme theory-building workshop was piloted with staff from FYD’s National Support Office to assess the utility and feasibility of a workshop approach. The pilot prompted revisions to the process which resulted in using a focus group interview format with the Project K staff and founders. A brief overview of the pilot process is provided below.

Workshop Participants

Eleven female and two male National Support Office staff members from various FYD departments including programme development, fundraising, marketing, communications, office administration, and national programme management took part in the hour long pilot.

Workshop Process

Attendees worked in small groups to generate content for the logic model components. Then, as a larger group, we attempted to conduct a rough thematic analysis and sequence the programme processes with relevant outcomes. The pilot workshop highlighted several problems with the workshop process (i.e. the length of time required was far greater than the time we had been allocated, not all groups wrote the outcomes of their discussions in sufficient detail, and participants needed a great deal of explanation to understand what was required). As a result, I decided to hold informal meetings in each of the regions and discuss the programme in a format more akin to a semi-structured focus group interview when approaching the Project K staff and founders. I felt this would make it easier to draw out the detailed responses needed to uncover the underlying programme processes and it would allow me to guide the questioning process more directly and clarify any confusion as soon as it arose.
Primary Data Source: Focus Groups with Project K Staff and Founders

Focus Group Participants and Data Collection Process

In earlier efforts to build positive relationships with the Project K staff (i.e. the consultative efforts discussed in Chapter Two), I identified programme directors who were interested in being involved in the evaluation. Seven of the eight agreed to participate in a discussion about the programme’s underlying theory. They were informed that other staff or board members were also welcome to attend the meeting, thus I also used a snowball sampling strategy. A total of 15 staff members (originating from seven different programme regions) took part in four focus groups consisting of between two and five people, and one one-to-one interview. An additional focus group was conducted with the two Project K founders. I kept detailed notes throughout the interview process. The comments made by the focus group participants were approximated as closely as possible to what was actually said.

The Focus Group Interview Schedule

As stated, the questions used in the theory development interviewing schedule were informed by Gugiu and Rodriguez-Campos’ (2007) SSIP (the general outline of the schedule is available in Appendix B). Below, I elaborate on the components and questions asked.

a) The antecedent condition: Renger and Titcomb (2002) argued that to understand a programme’s rationale one must first begin with understanding the situation that prompted the need for the programme. Hence, I asked: What situation in your community is Project K designed to resolve? or Which characteristics of the current social climate necessitate a programme like Project K?

b) The student profile: The target population should be identified by the social need described in the antecedent condition but it is also important to evaluate whether the programme actually reaches this population. I asked focus group participants to describe the typical PK student profile or the common factors associated with the group of students selected for the programme.
c) *The essential programme strategies:* To identify what was believed to be the key features of the programme leading to the desired outcomes I elicited discussions about the most important programme processes. In approaching this question I asked: *What are the most important processes occurring in the programme to produce desired outcomes?*

d) *The factors that influence the operation and success of the programme:* Keeping with Weiss’ (1997a, 1998a) theory of change approach, I also chose to include influencing factors or contextual influences. I asked: What are the key factors that would make this process likely to lead to a positive result or would impede it from happening?

e) *The key outcomes:* To identify the outcomes I asked: *What are the most important outcomes resulting from a student’s participation in the programme?* Outcomes were also explicitly drawn out by asking the focus group participants to track a causal sequence from a key programme process to its resulting outcome. I attempted to distinguish between short-term, intermediate, and long-term outcomes by asking the participants to continue with the sequence of tracking subsequent outcomes: *So if [process] leads to [immediate outcome], what follows from this? What does the [immediate outcome] lead to?* The topic of long-term outcomes was often broached separately as it was difficult to keep the interviewees on track with sequencing more than a few shorter-term outcomes.

**The Analytical Strategy**

As indicated in Figure 3, thematic analysis (as described by Braun & Clarke, 2006) was used to generate the initial themes which were then used to construct the programme theory. This analytical strategy allowed me to develop a holistic yet detailed picture of the programme; it also has the potential to illuminate unanticipated insights.

Since the initial step involved *generating the theory,* it made sense to first use an inductive approach where questions about the programme guided the analyses but no preconceived themes were formulated prior to beginning the analyses. To
do this, I followed the guidelines proposed by Braun & Clarke (2006). To summarise: I first familiarised myself with the data by reading and re-reading the focus group notes; I identified extracts that highlighted similarities and/or distinctions across the notes and labelled these with initial ideas for themes; any initial themes that represented similar constructs were then grouped into the same overarching thematic category; I reviewed the extracts associated with each thematic category and revised any that I felt were miscategorised; I then reflected on the core meaning of each thematic category and selected a final thematic label; finally, I selected representative extracts to provide evidence of the themes. The steps were repeated several times until I was satisfied with the final product, as Braun and Clarke recommend.

The sample of staff members included a group of individuals who differed greatly in their level of experience with the programme. As a result, the validity of a theme was evaluated by balancing both its prevalence across the groups and the expertise of the individuals producing the comments. It is important to note that while I did attempt to elicit causal sequences from the focus group participants when discussing the programme processes, I created the thematic categories and was responsible for interpreting how these themes could fit together to create a coherent programme theory.

**PRELIMINARY RESULTS**

Figure 4 presents the preliminary programme logic model. Each component is then explained in greater depth. The overarching themes and subthemes are represented in bold and accompanied by supporting quotes. For each relevant theme, the prevalence across the focus groups is shown alongside the theme in brackets. Participant quotes that are incorporated in the text are italicised and placed in quote marks. A brief overview of the model concludes the section.
Figure 4. Preliminary Project K logic model detailing the antecedent condition (purple), characteristics of the target group (blue), key programme processes (yellow), the influencing factors (light green and orange), short-term outcomes (pink) and intermediate outcomes (green).

THE ANTECEDENT CONDITION

Overall, the focus group participants expressed a belief that the current social climate is unsupportive to those at greater risk. This was articulated through three subthemes. First, the programme founders clarified that the impetus for the programme arose from negative health and wellbeing outcomes and the situation for youth in the country continues to be characterised by a high level of negative statistics (1).

[The initial idea for the programme was prompted by] the social recognition that New Zealand was leading the world in negative youth statistics... The issues haven't gone away - very little has changed. There seems to be even more poverty and deprivation. –Focus Group (FG) D

The second subtheme suggested that some young people are falling through the cracks. Many resources target youth at either end of the risk continuum, but there is a lack of support for young people on “the cusp of risk” (3).
Every one of us knows someone who is in the middle of the road and would benefit with some extra support. And education these days caters to the top and the bottom but not the middle. There’s a mass of kids that education is failing... There are so many kids and a lot of them get forgotten; programmes are either for high achievers or for those at the bottom of the cliff. –FG B

Unfortunately, two staff members also felt that there was a lack of community connectedness (2) for some of their youth participants and in certain localities. “The traditional networks that were there before aren’t as strong – the school and church.”

THE STUDENT PROFILE

It was clear the focus group participants understood the programme to be officially targeting youth with low self-efficacy (4).

It’s for those with self-efficacy but really for any kids who will most benefit.- FG D

One group did acknowledge that “predominantly it is low self-efficacy kids” that the programme actually reaches but also indicated that this feature was accompanied by other common characteristics. For example, the majority of their Project K participants had experienced some kind of disruptive life event (2). “Probably 60-70% have had changes in family foundations, maybe even more or they may have physically moved locations.”

Consistent across the majority of the groups was that there was a lack of “something” (4) that distinguished these students from those not eligible for the programme; however, the “something” that was lacking differed across the groups and, in general, “they are really diverse” (3). For example, one group mentioned a lack of family support and another spoke of poor decision making.

It’s the lack of family support. Not all, but the majority – like 90% have a lack of support or inappropriate or ineffective support. –FG E

They have a poor decision making process, they’re not conscious of their decision-making process. It’s an automatic process and they only have one perception of the world where this hasn’t been challenged. They’ve never been questioned. –FG F
THE ESSENTIAL PROGRAMME STRATEGIES

The focus group participants discussed numerous features of the programme that were thought to be critical to programme success. On analysis of the responses I felt that most related to an **experiential learning cycle** (3).

Two groups articulated experiential learning as a cyclical process that involved aspects such as feedback, processing or reflection, awareness and transference or repetition; they asserted that this cycle was intentionally taught by Project K facilitators.

All activities lead to self-reflection. The activities don’t matter; they are just a means to an end. It’s the experiential learning cycle and having the ability to take on feedback and apply it... Something happens (an activity or event) then you get information from it – both from others and internally – then you process it and reflect. You know, “I could do this that way next time or yeah, this worked well for me”, and then you apply it and talk about how to transfer it to other areas. –FG F

Another group expressed the process as “**doing activities that teaches them something**” and every group discussed at least two or more aspects of the experiential learning cycle mentioned above. For example, several of the groups mentioned that a key factor to Project K’s success is that it provides **new and challenging experiences** (5). Most groups then spoke about the importance of **feedback** (5), **reflection** (4), **awareness** (5), and/or **transference** (5) of the lessons learned. Next I describe each of these experiential learning subthemes in more depth.

**New and Challenging Experiences**

Table 2 summarises the different types of experiences the focus group participants discussed as being provided by the programme. It also shows the programme stage these experiences are associated with.
Table 2. *Types of New and Challenging Experiences Provided by Project K*

<table>
<thead>
<tr>
<th>Types of Activities/Experiences</th>
<th>Prevalence Across Focus Groups (n = 6)</th>
<th>Emphasis In</th>
<th>Examples of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teamwork</td>
<td>5</td>
<td>Wilderness Adventure and Community Challenge</td>
<td>Having to spend time with a group...and learning about working in a group, how to function with the group. You can't choose who you work with in life - FG B</td>
</tr>
<tr>
<td>Relationship Building</td>
<td>5</td>
<td>Mentoring</td>
<td>Maintaining a relationship outside of their family and social group and culture leads to respect and also learning their role in this relationship. It's about a relationship with someone who has a selfless investment. -FG C</td>
</tr>
<tr>
<td>Goal-setting</td>
<td>5</td>
<td>All Stages</td>
<td>Goal-setting - it’s a process. Thinking about “how to” -- how to make decisions. At this stage it’s usually low level goals but it’s getting them to understand the process. -FG E</td>
</tr>
<tr>
<td>Leadership</td>
<td>5</td>
<td>Wilderness Adventure and Community Challenge</td>
<td>I think leadership training is HUGE for some, not everyone but for some of them. They learn independence, how to do things themselves. They weren’t so good at this before. -FG E</td>
</tr>
<tr>
<td>Physical Tasks</td>
<td>4</td>
<td>Wilderness Adventure</td>
<td>Some challenges are harder for others. Like for some, getting in the abseil harness at the top of the cliff may be the hardest thing they’ve ever had to do. -FG A</td>
</tr>
<tr>
<td>Giving Back (The Community Project)</td>
<td>4</td>
<td>Community Challenge</td>
<td>The project leads to learning they can make a difference but also that if they don’t put in the effort it doesn’t make a difference. -FG C</td>
</tr>
<tr>
<td>Exploration</td>
<td>1</td>
<td>Community Challenge</td>
<td>Exploring the community connects them to the community so it’s real life. They learn about resources, start to set goals, and also a more positive vision of themselves. - FG D</td>
</tr>
</tbody>
</table>
Feedback

Project K was seen to underscore the importance of feedback by incorporating a group debrief following each learning experience. During the debrief facilitators acknowledge the participants’ individual strengths; they also provide constructive comments and solicit feedback from group members. One group explained that participants also gain feedback “internally from doing the activity”

The leader for the day gives a debrief of how the day went for him or her and then they feedback on others and the others feedback as well, as part of the debrief. It may all be positive or there could be constructive feedback as well. –FG A

Reflection and Awareness

Three groups explained that Project K provides “opportunities to think” and encourages reflection about the lessons the young people have learned, their future, or “how they’d like to change” and this generates a sense of awareness. Hence awareness may be considered an immediate outcome of involvement in the experiential learning process and this awareness may be developed in relation to themselves, others and/or their environment. I equated “awareness” with comments made in regards to knowledge development, realisation, understanding and learning.

Awareness of how the community sees them. Initially it may be negative but when they engage in activities it gives them positive publicity and they are aware of how they are viewed and then they might believe that they are in control of this themselves… They’re often surprised because people are helpful and they didn’t realise they could just walk up to people and ask and then they would help. They start realising resources and opportunities. –FG F

Transference

The learning cycle ends and begins again when participants transfer their awareness and learning to a new setting. Project K facilitates this transference by drawing comparisons between previous experiences and ones that may arise in the participants’ day to day situations, and by emphasising how what was learned previously can be applied to a new situation.
There’s transference to a variety of activities, like especially with things not being such a big deal anymore, for example jumping out of a helicopter. -FG E

As explained by several of the focus group participants, the three programme components (the Wilderness Adventure, the Community Challenge, and Mentoring) simply provide different contexts for the same experiential learning process to occur and the repeated emphasis of this process to participants serves to cement the learning.

FACTORS INFLUENCING PROGRAMME SUCCESS

Although the focus group participants were asked directly about the factors that increase or decrease the likelihood of a successful programme, sustaining programme outcomes was often discussed as being contingent on other factors. Participants discussed these concepts using “if…then” and “it depends on…” statements. This signalled that these processes were important influencing variables. Using both types of information, I identified four overarching categories of moderators: the Project K environment, individual characteristics, interpersonal dynamics, and ongoing input. Participant engagement was also identified as a vital ingredient to programme success but because of its central role in the psychological change process of experiential learning its position differed to the other influencing factors. This is discussed in greater detail below.

The Project K Environment

The Project K environment is purposefully constructed to be an intense but safe and autonomy supportive environment and this combination appears to generate an optimal level of arousal and engagement in the learning process. The environment is intense because the level of challenge is sufficient, the intimacy of the Project K group and the “foreign environment” magnify the experiences, the youth participants are pressured to persevere, and there is limited choice.

Having at least a few activities that really get them out of their comfort zone – this really brings out differences in strengths and weaknesses that people are unaware of. -FG A
Because it’s intimate they can see the effects right away and everything is used as a learning opportunity. The whole situation is magnified. –FG A

When they are away the options are taken away from them...During the Wilderness Adventure they can’t not tramp. They can’t give up. –FG F

It’s about perseverance because they have no choice but to do it. –FG D

However, one person noted that “it’s important to distinguish what is outside of their comfort zone with what is frightening. It shouldn’t be frightening”. Related to this is the subtheme of a safe environment. Several staff members felt that Project K participants would attempt to try new things and take on challenges within the Project K environment because they could trust that their physical and emotional safety would be preserved. This includes being in “an environment where it’s Ok to make mistakes...”

The Project K environment was seen as supporting the participants’ autonomy through leadership training, a student-driven community project, having the participants explore their own career interests and decide their own goals.

Ownership of the challenge is important and the more input they have, the more ownership. If it’s their own then there’s more success...Keeping it fun is really important--if it’s creative. This is related to ownership. If it has a life to itself they want to be part of it. –FG F

Facilitators must, however, strike a balance between a student-driven approach and pointing the young people in the right direction to keep them on track, as one individual mentioned:

Give lots of help at the start then facilitation starts to move away to let them take over. You can’t do this at the beginning but it is important. –FG F

**Individual Characteristics**

The focus group participants spoke of three different types of individuals that can impact on programme success: the programme facilitators, the mentors, and the youth participants. Their *individual characteristics* include skills and personality traits as well as the personal environment or social histories that these individuals carry with them.
Being able to create an environment that is responsive to individual characteristics requires skill and self-reflection. As noted above, facilitators must also be able to scaffold learning in a manner that promotes participant autonomy within some broad guidelines and be available for additional support when needed. Not surprisingly, the significance of the facilitators’ qualities (6) was acknowledged in every group.

It’s the experiential cycle and how it’s delivered - the facilitation, the staff on the ground. We know from experiences how much better it can be with better people since we changed some. –FG E

In a similar vein, individuals in the majority of the focus groups stressed that mentor characteristics (5) also exert a major influence on the participants’ perspectives of the programme.

At the end you hear the different comments: some say it was amazing, others say “I wish my mentor was” this or this or this. They can just tell it would have been so great if they got that other really great mentor. –FG E

The varying personalities and previous personal experiences of the Project K participants (4) play an important role in their individual level of engagement in programme processes and/or growth during and after the programme, according to participants in four of the groups. One staff member described four broad types of Project K participants she had observed during her experiences with the Project K and she outlined how these influenced their success in the programme:

There are three no maybe four types of students: some run with it and it’s life-changing…; others just go through the motions just because they have to; others aren’t fussed during the mentoring but at the end they look back and wish they had put more into it once they realise what they could have got out of it; and others are just disengaged and don’t ever get into it – they don’t take part. Their mentor may contact them and then wait weeks trying to see if they’ll take the initiative and then they get encouraged to keep trying with support and eventually they may scrape through or not finish. –FG A

Interpersonal Dynamics

The characteristics of all the individuals involved with the programme determine the interpersonal dynamics and these also have important implications for programme success. The individual who provided the above quote also explained how a young person’s preferences and openness to new experiences
determines their responsivity towards different types of mentors and if this is not taken into consideration then the **mentor-mentee match (5)** may not work out. A young person’s needs also appear to be a central consideration when determining which match will work best.

> Matches are critical...matching with needs is important. If they are both looking to hang out or want a brother then good but if others just want information and already have a family, a mentor who wants a little brother is not as good a match. –FG F

Another dynamic that was salient to the majority of the focus group participants was that of the **Project K peer group (5)**. According to several staff, the way a Project K team works together can largely influence the success of the group-based programme components. Luckily, the programme personnel can play a role in how well a group works together by being aware of and facilitating this dynamic.

> The group dynamic is also important and what’s going on with them as individuals. Facilitating the group dynamic is a big part – addressing any problems. –FG F

**Ongoing Input**

Ongoing input, as conceptualised here, incorporates both interpersonal support and programme infrastructure aspects that are used to amplify programme effects and help the young person sustain positive changes.

*Interpersonal support*

Interpersonal support is a theme that permeated all of the discussions. **Mentor support (6)** is the explicit aim of the final programme component and this was affirmed in all of the focus groups.

> The mentor takes them through smaller progression steps. For example, if they need to have a conversation with mom, they can first try it with the mentor. - FG F

The **Project K peer group (3)** also provides an important source of support especially during the first two components of the programme. A staff member explained how “in mentoring they have to break up and it’s hard cause they drew so much support from that [PK] group.” Support in the form of **positive reinforcement (4)** also occurs via feedback from mentors, staff, PK peers, and the community. One
of the programme founders explained that “celebration is an important part of this. It reinforces the buzz and the accomplishment with the group”.

The multiple sources of support ensures that it is frequent and ongoing for the full 14-months of the programme but the importance of interpersonal support was also reiterated when discussing what happens once the programme is complete; if the youth participants have a supportive external social network (4), especially in the form of family then sustained outcomes are much more likely, as reported by many of the focus group participants.

I don’t know what happens once it’s [the programme] gone- if it sticks. It depends on if they have support to help them through the next bit. –FG E

Family support can also influence the ability Project K has to influence the young person during the programme, as another participant pointed out:

The mentor may work with the family or they may not but then it may not be as successful…Parent agendas can be hard. For example, we had someone in a… family paired with a high profile mentor and the family was terrified and wouldn’t let it happen, they wouldn’t let the mentor come to the house. We have to abide to parent agendas. –FG D

Programme infrastructure aspects

Project K’s infrastructure can also contribute to sustained positive growth. Specifically, Project K is purposely designed as a staged progression of services (3), to reinforce the experiential learning that occurred in the previous stage, as was alluded to earlier, and to ensure that support occurs over a lengthy period.

Debriefing is important to process things then it leads to transference, but they would lose this if they don’t do other components of the programme because they’re back in their reality and it’s still shitty. –FG D

The Wilderness Adventure may spark a desire or success that is maintained during the Community Challenge. The mentor relationship keeps the flame alive. –FG F

Participant Engagement

Several focus group participants pointed out that participant engagement (6) is critical to the learning process, but this was primarily discussed in relation to the Community Challenge. As one person put it, “they need to want to and engage
It was evident that the difficulty staff had in engaging the participants during this programme stage made this factor more salient.

Giving back is important but sometimes it just seems like a gesture, like they’re ticking a box... if it’s not a gesture, it does give them a sense of achievement – they remember it, how they helped. They remember the kid’s smiling faces when they helped at the home, etc... If they don’t see why they are doing it. They ask “Why do we have to?”... seeing the link of its importance is important, but it’s not like we don’t tell them or teach them the link. You can’t teach it, they just don’t process it. – FG E

Related comments included the need for participants to understand the meaning of the activities and the importance of making the activities fun. Here the community challenge is contrasted with the Wilderness Adventure:

[The Community Challenge is] boring or not as exciting as the Wilderness Adventure– there’s not that adrenaline side of it and they tend to lose motivation with it... It doesn’t feel like it has the same impact. – FG A

When constructing the logic model I felt it made intuitive sense to place participant engagement at the centre of the experiential learning cycle because without cognitive or behavioural engagement it seemed unlikely that a participant would process the feedback provided and, consequently, develop a new sense of awareness. Yet I also wanted to isolate this construct because it was not discussed as part of the experiential learning cycle by the interviewees and was not a programme strategy per se.

**THE KEY PROJECT K OUTCOMES**

**Short-Term Outcomes**

Table 3 provides a breakdown of the short-term outcomes, their prevalence across the focus groups, common subthemes falling under each outcome category, and examples of comments supporting these themes. It was obvious to each group that the programme provided opportunities that facilitated the development of participant skills and awareness of themselves, others, and the social or physical environment. Every group identified one or more outcomes associated with a **positive self-concept** (such as confidence, self-esteem, and/or agency), **skill development**, and **connectedness**. In terms of skill development, the most common were social and goal-setting skills but organisational, problem-solving...
and leadership skills were also mentioned. The focus group participants discussed connectedness-related outcomes with respect to developing new relationships, a sense of community, and greater trust in adults. All but one group talked about participants being able to **recognise opportunities and resources** as a result of the ideas seeded by Project K and the doors it had opened for them. **Social conscience** was also a frequently reported outcome category (i.e. valuing others, their communities, and respecting the environment). Less commonly reported outcomes included aspects reflective of **independence** such as self-sufficiency and accountability; **affect-related outcomes** (e.g. enjoyment); **a sense of achievement**; and **positive recognition** from the community or school.
### Table 3. Immediate and Short-Term Outcomes with Associated Quotes

<table>
<thead>
<tr>
<th>Short-Term Outcome</th>
<th>Prevalence Across Focus Groups (n = 6)</th>
<th>Subtheme</th>
<th>Examples of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Self-Concept</td>
<td>6</td>
<td>Self-Esteem</td>
<td>...the feelings of self-worth – that someone else wants to invest time and energy in them for nothing and they are worthy of this investment and it’s unconditional. - FG A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Efficacy</td>
<td>They gain confidence – in communicating, in their abilities, and in trying new things. -FG C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agency</td>
<td>When it comes to PK they get a sense of power and ownership over their lives. - FG F</td>
</tr>
<tr>
<td>Skills</td>
<td>6</td>
<td>Social</td>
<td>They also mix with students that are not their best friends or even people they don’t like и they have to work through this to get through the tasks and focus on their communication skills… - FG A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Goal-setting</td>
<td>Goal-setting – it’s picking some goals and knowing the process of how to get there so they can achieve their potential. It’s a positive habit forming process. - FG D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Organisational</td>
<td>The organisation and planning to help with study, either the mentor prompting them or them setting a goal around this with the mentor showing them is important. -FG A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Problem-Solving</td>
<td>Also they are held accountable and taught skills to solve problems to overcome obstacles. -FG F</td>
</tr>
<tr>
<td>Connectedness</td>
<td>6</td>
<td>New Relationships</td>
<td>A huge strength of this is that they shared this amazing thing in the Wilderness Adventure and they can’t share it with anyone else. They have this special bond with this group and it’s still as tight as when you see them later. You hear them when they get together again, chatting about remember this and that, etc. It’s really amazing. They shared something. - FG B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sense of Community</td>
<td>Exploring the community – connects them to the community so it’s real life. They learn about resources… - FG D</td>
</tr>
</tbody>
</table>
## Project K’s Theory of Change

<table>
<thead>
<tr>
<th>Short-Term Outcome</th>
<th>Prevalence Across Focus Groups (n = 6)</th>
<th>Subtheme</th>
<th>Examples of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Trust</td>
<td>Also learning to trust another adult/a stranger, someone who will do it for nothing, and having (we hope) a decent and reliable adult in their lives. This creates awareness that there are decent people out there. That they can choose the people they have in their life.- FG F</td>
</tr>
<tr>
<td>Recognising Opportunities &amp; Resources</td>
<td>5</td>
<td></td>
<td>They’re often surprised because people are helpful and they didn’t realise they could just walk up to people and ask and then they would help. They start realising resources and opportunities.- FG F</td>
</tr>
<tr>
<td>Social Conscience</td>
<td>5</td>
<td></td>
<td>They also have to mix with students who are not their best friends or even people they don’t like and they have to work through this to get through the tasks and focus on their communication skills. This seems to lead to them valuing others and valuing differences.- FG A</td>
</tr>
<tr>
<td>Independence</td>
<td>3</td>
<td></td>
<td>They realise they are self-sufficient..They are so removed from their own group, they can only rely on themselves.- FG B</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>3</td>
<td></td>
<td>An experience of feeling good and successful where they are not used to this in their home environment. -FG F</td>
</tr>
<tr>
<td>Achievement</td>
<td>2</td>
<td></td>
<td>Giving does give them a sense of achievement. They remember it, how they helped...- FG E</td>
</tr>
<tr>
<td>Positive Recognition</td>
<td>2</td>
<td></td>
<td>It’s about knowing they had an experience that is unique and valued by the community. -FG F</td>
</tr>
</tbody>
</table>
Intermediate Outcomes

The intermediate outcomes were those the focus group participants discussed in reference to what they observed in individuals towards the end of the programme or soon after the programme had finished. The most prevalent intermediate outcomes pertained to increased motivation (4) such as perseverance and resilience. One person described how the participants would “meet obstacles but keep going – they look at new options” and another explained how Project K cultivated “resilience when things don’t go right”. Related to this were comments about initiative. After the programme, Project K participants have the “ability to say ’yes’ when opportunities arise and stick it through”.

Half of the groups spoke about participant resourcefulness or help-seeking (3) behaviour and the fact that participants were “better communicators” (3) after having experienced the programme.

The also seek environments now that would support them. For example, they think about “What school would be the best for me?” –FG F

...because of PK they are more likely to ask for help, because of sharing and getting positive feedback from it.- FG D

Long-Term Outcomes

There was much less conviction when discussing longer-term outcomes. The focus group participants spoke about how “some people are…” and how “anecdotes about participants later on revealed….” Some admitted they could only presume or hope for long-term outcomes because they often did not maintain contact long after the programme. Four of the groups envisioned long term outcomes related to a positive future direction (4). For example, they hoped participants would “stay in school”, be “career or goal-oriented”, or having “better employment opportunities”. Two groups talked in a similar way about leadership (2).

Anecdotes reveal that one year post-programme they are starting to become leaders. We get stories of this; they become public speakers when they never had any intention to. We hear of some running toastmaster groups. –FG D
A couple of the groups also indicated that the outcomes were “student dependent” (2). The following individual felt strongly that specifying concrete outcomes was a problem:

You look at the role development for each of them—individual growth relative to their own beginning. You need to value this and this is missed with blanket outcomes. -FG C

Others mentioned one-off examples of what they envisioned for future participants, for instance “becoming better parents”, being in “stable relationships”, or being of “enhanced economic value” to New Zealand society.

The preliminary programme logic model is summarised below. Because there was such uncertainty and a lack of consistency in the way that the interviewees discussed long-term outcomes, I opted to exclude these.

**SUMMARY OF THE STAKEHOLDER-DRIVEN LOGIC MODEL**

Because the current social climate is perceived to be one that is not supportive to young people at higher than average but not extreme risk, the Project K programme was developed to provide support to 14-15 year old Year 10 students with low self-efficacy. In reality, the participants may have low self-efficacy, but also seem to lack “something” else, and in many cases have experienced disruptive events in their lives. Still, the programme accommodates a fairly diverse set of young people.

The three components of the programme provide different opportunities to take part in new and challenging experiences which set in motion the experiential learning cycle. Programme personnel, such as the Wilderness and Community facilitators and mentors are equipped with valuable skills and guide participants through this process by providing explicit feedback and soliciting feedback from their Project K peer group, encouraging self-reflection, and drawing comparisons between the lessons learned during the programme to other areas of the participants’ lives. The intention is to raise participant awareness and encourage the young people to apply their knowledge and skills to new experiences in the future. The nature of the Project K environment, ongoing input from the
programme personnel and the Project K group, and the continuation of programme services over a 14 month period help keep the young person engaged in the learning process and reinforce the lessons learnt but this also depends on the characteristics of the people involved, how these individuals work together and additional support received from external social networks.

If the young person remains engaged throughout the process, immediate outcomes relating to a more positive self-concept, knowledge and skill acquisition, a feeling of connectedness with others, the recognition of resources and new opportunities, and social conscience are likely to occur. Other immediate outcomes may include independence and self-sufficiency, a sense of achievement, positive affect, and positive recognition from others. Later outcomes may include increased motivation, resourcefulness and help-seeking behaviour and effective communication, in addition to sustaining the more proximal outcomes.

**STEP 2 OF EXPLORATORY THEORY DEVELOPMENT: ASSESSING CONSENSUS**

**THE ANALYTIC STRATEGY**

As indicated earlier, three comparative data sources were analysed to assess the consensus of the findings from the preliminary model: Project K youth participant responses to an end of programme survey; four key programme documents; and six previously-conducted Project K research studies. Initially, I used the inductive thematic analysis approach detailed above to analyse the Project K participant survey comments: Though the themes identified in the semi-structured interviews were kept in mind, I reviewed the comments from these two sources with an open mind to new themes. The initial results from the youth participant sample and the findings from the programme document review and the previous Project K research studies were then evaluated against the themes identified in the focus groups with the programme staff. The findings from each are presented next.
COMPARATIVE DATA SOURCE 1: YOUTH PARTICIPANT RESPONSES

As the first step in assessing the consensus of the preliminary programme theory across multiple programme sources, I reviewed the open-ended responses to the question “What did you gain from participating in Project K?” posed to Project K participants in an end of programme survey developed by FYD. Although the question could not yield information that would be directly comparable to the focus group results about the antecedent condition, student profile, or moderating influences, I felt the responses from the youth participants would allow a good comparison of the outcome themes and the processes in so far as they are proximal outcomes.

Project K Respondents

The respondent sample involved in this study came from 40 different programme sites that were included in Project K’s randomised controlled trial outcome evaluation (described in full in Chapter Five)\textsuperscript{12}. These programmes were delivered between the end of 2004 and the end of 2006 across 11 regions of the country. At the beginning of the programme, the 40 programmes included 476 Project K participants; however 65 did not graduate from the programme leaving 411 graduates available to complete the survey at the end of the programme. Of these eligible Project K graduates, 351 responded to the question, representing 85% of the full sample. Females accounted for 47.3% (166) of these responses.

The survey allowed participants to select more than one ethnicity and the Total Response method of coding ethnicity, where participants are counted once for each category selected was utilised to calculate participant frequencies within each ethnic category (more details on Total Response coding are provided in Chapter Five or see Allan, 2001). In consequence, one individual could be counted more than once. Ethnicities were grouped into the 5 most common New Zealand ethnic groups: The European category (both originating from New Zealand and

\textsuperscript{12} The full randomised controlled trial included 50 different programme sites but the programmes delivered in 2007 were not included in this study sample because the data were not available at the time of analysis.
Project K’s Theory of Change

those outside) contained 280 participants; the NZ Māori category had 82 participants; the Pacific peoples category (e.g. Samoan, Tongan) included 43 participants; the Asian category (e.g. Chinese, Indian) contained 14 individuals; and 5 individuals selected the Other category (for those of identifying with an ethnicity not listed).

Youth Participant Results

In the eyes of the programme participants, Project K was effective in producing a variety of positive outcomes; 99% of these participants reported at least one positive gain. These gains were organised into 9 main themes: a) Self-concept, b) Knowledge and Skill Acquisition, c) Striving & Achievement, d) Connectedness, e) New Experiences, f) Positive Outlook/Attitude Change, g) Affect-Related Changes h) Physical Health, and i) Maturity.

Table 4 presents each of the 9 themes, the prevalence of the theme across the 351 respondents, and the subthemes within each overarching outcome theme. The prevalence of a subtheme is indicated if it was reported by a substantial number of individuals. The table demonstrates that the most prevalent theme related to a change in self-concept (mentioned by 63% of individuals). Respondents indicated that they were more aware of themselves, they had higher self-esteem, and they recognised that “...it’s ok to not be like everyone else”, but the confidence subtheme was the most dominant within this category with 56% of the total sample noting this particular outcome. Knowledge acquisition and skill development was the next most popular category with 46% of respondents reporting gains primarily in relation to goal-setting (20%) and social skills (15%), but also in leadership, outdoor, problem-solving, and general life skills. Twenty percent of the respondents also reported improvement related to aspects of motivation or striving such as persevering through challenges, being more focused, and having the initiative to participate in activities. Related to this, some also identified achievements they had obtained. This was followed by a theme of connectedness. Connectedness was usually expressed as the development of new friendships or the availability and receipt of support. All other outcome themes were mentioned by fewer than 10% of the youth participants. Nevertheless, several of these (i.e.
affect-related outcomes, maturity, and positive outlook) fit with some of the more minor themes noted by the staff focus group participants. A small portion of the youth participants (5%) also identified a gain related to physical health but this was not a salient outcome according to the focus group participants.
Table 4. Outcomes identified by Project K participants

<table>
<thead>
<tr>
<th>Overarching Outcome Theme</th>
<th>Prevalence</th>
<th>Subtheme</th>
<th>Examples of Comments</th>
</tr>
</thead>
</table>
| Self-Concept                | 63%        | Confidence (56%), Self-Esteem, Self-Efficacy, Self-Awareness | knowing it’s ok to not be like everyone else  
I have a lot more confidence to ask for help, speak up and I make new friends easier. After the wilderness challenge I realised that I am capable at physical activity and that boosted my self-esteem. |
| Knowledge/Skill Acquisition | 46%        | Goal-setting (20%), Social (15%), Organisational, Self-Regulation, Outdoor, Life, Leadership | The skills to talk to people.  
How to set goals, time management.  
A lot of knowledge about New Zealand and Upper Hutt and making and achieving goals. |
| Striving/Achievement        | 20%        | Motivation, Perseverance, Determination, Focus, Involvement, Goal Achievement | Motivation to do things and getting into activities.  
I have an idea of where I want to go in the future.  
To do my best in school and there more in life than just the playstation |
| Connectedness/Friends       | 19%        | New Relationships, Trust, Support                      | The ability to trust people and talk to people among my age group.  
I’ve basically gaved a new family, my mentors family and we are really close and we’ve been there for each other in the past year. |
| Positive Outlook            | 8%         | New Perspectives, Tolerance for Others, Attitude Change | To be positive and smile.  
a better outlook on the different types of people |
| New Experiences             | 6%         | New Activities, New Interests                         | I have done quite a few things that I would never have done.  
…discovering new things like New music |
| Affect-Related Outcomes     | 5%         | Enjoyment, Happiness, Emotional Regulation             | I had fun on the adventure programme  
Less afraid of heights  
Ability to control tantries. |

13 Original wording and spelling retained.
<table>
<thead>
<tr>
<th>Overarching Outcome Theme</th>
<th>Prevalence</th>
<th>Subtheme</th>
<th>Examples of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Health</td>
<td>5%</td>
<td>Fitness, Physical Activity, Healthy Eating</td>
<td>Am now much fitter. eat more variety of food no longer anti-social, don't drink, smoke, etc.</td>
</tr>
<tr>
<td>Maturity</td>
<td>2%</td>
<td>Independence, Responsibility</td>
<td>I feel I have matured a lot and I feel I have considered lots of options for my future. im way more independent and mature about everyday things</td>
</tr>
</tbody>
</table>
Alignment with the Essential Programme Strategies

With respect to essential programme processes or strategies, we can see that some of the outcome themes included in the above table align with aspects of the experiential learning cycle components identified in the staff focus groups. This is not surprising because, as mentioned above, some of these can be considered to be proximal outcomes. Approximately 6% of the Project K participants considered the new experiences provided by the programme as a salient gain. It was more difficult to find any references to obtaining feedback or reflective thought processing; however, one youth participant claimed, “I’ve learnt more about myself and from talking to others that I don’t know” and two others noted that Project K helped them to think about their future and to consider their options more carefully. Awareness was also a central feature of the experiential learning cycle described by programme staff and many youth participants recognised self-awareness as a programme outcome, though in the above table it is included as a subtheme in the overarching self-concept outcome category. Similarly, many youth participants noted knowledge acquisition as an important outcome and this corresponds with the conceptualisation of awareness described in the focus group results. Numerous Project K participants commented on their ability to exercise new skills in their current day to day circumstances which is suggestive of a process of transference. In addition, three participants explicitly articulated that the skills they had obtained on the programme were transferable to their normative environment. For example, one participant stated “I have gained many valuable life skills that I can apply to everyday situations, including setting goals, socially, in work and other areas”.

COMPARATIVE DATA SOURCE 2: DOCUMENT REVIEW

Four key documents were obtained from FYD’s Evaluation Coordinator as sources for the document review: The Project K Policies and Procedures Manual; the Project K Evaluation Manual; the Project K Liaison Manual; and the Project K Facilitator Manual. The Policies and Procedures Manual outlines the values and philosophy of Project K as well as the policies and procedures the regional Project
K staff and board members are required to meet under the licensing agreement. The Evaluation Manual outlines the evaluation processes and procedures and all required documentation for implementing both formative and summative evaluative activities. This includes the evaluation measures; information sheets for the school, teachers, caregivers, and students involved, and the associated consent forms. The Liaison Manual delineates the process for liaising with schools, students, caregivers, and mentors as well as the programme content that needs to be communicated to these groups, including detailed guidelines and documentation for monitoring the mentoring relationship. Finally, the Facilitator Manual provides direction and all of the content material for the Induction, Wilderness Adventure and Community Challenge, including all of the forms to be completed by the facilitators and students over the course of the programme. It also provides a general overview of the philosophy, values, policies and procedures of the programme.

A deductive thematic analysis was employed: Each of the documents was scanned for content that fit with the earlier themes identified in the Project K staff and founder interviews. Table 5 lists the themes and indicates whether they were included in each of the four documents. Moreover, if substantive evidence of a new theme not identified in the preliminary programme theory was apparent in any of the four documents, this was noted and these additional themes are discussed below.
Table 5. *Project K Document Review Findings*

<table>
<thead>
<tr>
<th>Logic Model Component &amp; Theme</th>
<th>Policies &amp; Procedures</th>
<th>Evaluation</th>
<th>Liaison</th>
<th>Facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANTECEDENT CONDITION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NZ youth population is characterised by negative statistics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Lack of community connectedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of support for youth on the cusp of risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TARGET GROUP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13-15 year olds with low self-efficacy</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>STUDENT PROFILE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Youth who lack “something”</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Youth who have experienced disruptive events</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A diverse group of young people</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>ESSENTIAL PROGRAMME STRATEGIES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiential Learning</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>New &amp; Challenging Experiences</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Feedback</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Reflection &amp; Awareness</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transference</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TYPES OF ACTIVITIES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploration (of community and resources)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Physical Tasks</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
## Logic Model Component & Theme

<table>
<thead>
<tr>
<th>Logic Model Component &amp; Theme</th>
<th>Policies &amp; Procedures</th>
<th>Evaluation</th>
<th>Liaison</th>
<th>Facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal-setting</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Giving Back (Community Project)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teamwork</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relationship Building</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## INFLUENCING FACTORS

<table>
<thead>
<tr>
<th>INFLUENCING FACTORS</th>
<th>Policies &amp; Procedures</th>
<th>Evaluation</th>
<th>Liaison</th>
<th>Facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant Engagement</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intense Environment</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safe Environment</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomy Supportive Environment</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facilitator Characteristics</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mentor Characteristics</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant Characteristics</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mentor-Mentee Match</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Dynamics</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mentor Support</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project K Peer Support</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staged Programme Services</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive Reinforcement</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External Support</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## SHORT-TERM OUTCOMES

<table>
<thead>
<tr>
<th>SHORT-TERM OUTCOMES</th>
<th>Policies &amp; Procedures</th>
<th>Evaluation</th>
<th>Liaison</th>
<th>Facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Self-Concept</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Skills</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Logic Model Component &amp; Theme</td>
<td>Policies &amp; Procedures</td>
<td>Evaluation</td>
<td>Liaison</td>
<td>Facilitator</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------------------</td>
<td>------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Connectedness</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognising Opportunities &amp; Resources</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Conscience</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Independence/Self-Reliance</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Positive Affect</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Achievement</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Positive Recognition</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**INTERMEDIATE OUTCOMES**

<table>
<thead>
<tr>
<th></th>
<th>Policies &amp; Procedures</th>
<th>Evaluation</th>
<th>Liaison</th>
<th>Facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resourcefulness &amp; Help-Seeking</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective Communication</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Alignment with the Preliminary Logic Model Data Source

The results demonstrate that each theme or subtheme noted by the Project K staff and founders was included in at least one of the programme documents, with the exception of some of the \textit{antecedent condition} and \textit{student profile} themes. Specifically, only the theme of negative youth statistics was included as part of the rationale for the programme, and participants were not described as potentially having experienced disruptive life events.

Although the long-term outcome themes were not included in the table for reasons outlined earlier, the focus of the programme documents on vocational direction and career-decision self-efficacy outcomes in a way reflect the desire expressed in the focus groups that the programme point youth participants in a positive future direction.

Discrepancies with the Primary Logic Model Data Source

As with the youth participant data, a theme that was palpable across each of the programme documents but was not discussed by the focus group participants was the desired outcome of improved attitudes and behaviour regarding health and fitness. This was a definitive programme objective as outlined in each of the four documents and a substantial portion of the programme curriculum was dedicated to aspects of physical health. Furthermore, the concept of attitudinal change was not particular to the domain of health and fitness; a strong emphasis on promoting positive attitudes in general was apparent in the Facilitator Manual. Actually, the theme of positive attitude promotion was not entirely absent from the focus group data; it was simply not articulated as explicitly as the content in the programme documents. Indeed, the social conscience outcome theme identified in the focus group feedback incorporates ideas such as respect for the community and valuing others; these ideas are essentially attitudinal. It could be said that positive changes in self-concept reflect changes in attitudes towards the self. It is also of note that attitude change or a more positive outlook was a gain reported by 8\% of the Project K participants.
A second notable theme that was more or less absent from the semi-structured interview data was cultural consideration. The policies and procedures clearly stated the need for culturally appropriate processes and sensitivity to the participants’ different ethnic backgrounds. Moreover, the stated learning outcomes included developing awareness about and honouring one’s own culture and respecting that of others. Some of the staff focus group comments did describe an increased tolerance and respect for diverse viewpoints on the part of the participants which implies a sense of cultural respect but this was not specified.

Finally, all of the programme documents outlined the process for excluding unsuitable students from the participant selection pool. Understandably, when asked about the typical student profile the Project K staff and founders did not discuss what characteristics the students did not have\(^{14}\), but the exclusion process was developed to clearly delimit aspects of the typical student profile and thus contributes to our understanding of this logic model component. Students who exhibit characteristics of violent behaviour, substance abuse, severe cognitive or learning difficulties; who have attempted suicide; or who are identified by a school guidance counsellor as not being suitable for the programme are excluded from participation.

**Unintended Side Effects**

Because an explicit question about unintended side effects was not included in the interview schedule, the issue was largely absent from the interviews (I noted in retrospect that this was a limitation of my interview process). That said, acknowledging the potential for such unintended negative effects is an important part of the theory-driven evaluation process; thus, the data sources were reviewed with an eye for any content that would be suggestive of this. Individuals from two of the Project K staff focus groups mentioned one-off comments that suggested potential contributors of negative effects: one person described how mentors that do not fulfil their year-long commitment can leave their mentees “gutted”; another

---

\(^{14}\) One focus group did discuss how the variability in the way the exclusion criteria are applied at different programme sites influences the kind of students that are selected. This is one potential reason for the diversity across the programme groups.
individual from another group spoke about the difficulty students have when trying to run the Community Challenge during the school-year because of the time missed during both the Wilderness and Community stages. The latter comment relates to a concern relayed in one of the programme documents. One section of the Liaison Manual covered strategies to mitigate the difficulties programme participants may have when integrating back into their classrooms, home life, and communities following the Wilderness Adventure.

**COMPARATIVE DATA SOURCE 3: PREVIOUS PROJECT K RESEARCH**

Because FYD has an internal evaluation unit, welcomes external research and evaluation, and has strong partnerships with two universities, Project K has been the focus of several other evaluation research projects, as discussed earlier. As a final source for triangulating the preliminary programme theory findings, I reviewed the results from the six relevant Project K research projects outlined in Chapter One. The findings from each of these studies contribute to our understanding of one or more aspects of the logic model, and largely corroborate what was found in the other data sources. What is more, these studies did not point to any components that had been overlooked.

Somervell’s (2011) exploration of the process used by Project K administrators to exclude unsuitable students from the programme selection pool demonstrated that the programme does in fact reach its stated target group of Year 10 students with low self-efficacy. That said, the fact that the exclusion criteria are not applied in the same manner across the programme sites also suggests that substantial variation would exist in the characteristics of the Project K groups across the country. A few of her interviewees (programme directors and school personnel involved in the exclusion process) did note that the use of the participant selection survey can be beneficial in identifying young people who suffer silently as a result of internalising their distress and thus would normally get overlooked by other support service providers. This theme fits with the idea that support and resources are lacking for youth who are at higher but not at extreme risk. Her interviewees also reported that the purpose of the exclusion process is to ensure
the safety of the other participants and to ensure that the programme is a good fit for all individuals who are included. Thus the exclusion process in and of itself suggests something about the importance of a safe environment, participant characteristics, and the team’s dynamics.

The fact that the Māori graduates from Hollis et al.’s (2011) narrative interview study attributed some of the positive programme effects to the challenging and reflective activities supports these aspects of the experiential learning cycle. Their positive comments about the way they were supported by their instructors and mentors also substantiate the positive influence of an autonomy-supportive but scaffolded learning environment and of mentor support. Zhang’s (2011) investigation of Project K’s health and fitness aspects also pointed to experiential learning, reflection, mastery experiences, and goal-setting as critical programme processes, and external support, individual differences, and the manner the programme is delivered as key influencing factors. Warren’s (2005) interviews with Project K students before, during and after the programme also corroborate the importance of mentor support, the difficulty of engaging participants in the Community Challenge, and the development of goal-setting competencies and connectedness.

O’Neill’s (2005) study on the Project K mentor training programme demonstrated the value the organisation places on selecting highly skilled and knowledgeable mentors and it also hinted that developing awareness about the mentees’ cultural backgrounds is important. The focus on cultural awareness was not an explicit part of the preliminary model; however, it was identified as an important theme in the document review and Hollis et al.’s (2011) findings also suggest that a lack of cultural competence or sensitivity can create unnecessary stress for ethnic minorities.

Qiao and McNaught’s (2005) analysis of the randomised controlled trial outcome evaluation data for a subsample of Project K programmes demonstrated that the programme is successful in effecting positive changes related to the youth
participants’ self-concept, and that participant characteristics (e.g. ethnicity) can influence this process.

Returning to the findings from Zhang’s (2011) research, she showed that the programme can have a limited effect on influencing positive health behaviours, at least in the short-term. However, a focus on health and fitness, while an important aim, was recognised by staff to be secondary to cultivating a positive self-concept.

**DISCUSSION**

This theory-building process was beneficial for a number of reasons: it was a strategy for assessing the programme theory’s level of consensus across the organisation; it served as platform from which the legitimacy of the theory could be assessed against the academic literature; it provided some insight into programme implementation issues; it provided direction for future evaluative activities; and it contributed to knowledge about youth development processes and general programme evaluation practice. These contributions are explicated in greater detail below.

**MODEL FIT ACROSS THE PROJECT K DATA SOURCES**

The final version of the Project K Programme Logic Model is presented in Figure 5 and is the result of compiling the findings across all of the programme data sources. Because the evidence across the comparative data sources was overwhelmingly consistent with regards to the target group, programme strategies, influencing factors, and the majority of the outcome components identified in the development of the staff-driven logic model there was little need for revision.
Figure 5. The Finalised Project K Logic Model depicting the antecedent condition (purple), characteristics of the target group (blue), key programme processes (yellow), the influencing factors (light green and orange), short-term and intermediate outcomes (pink) and desired long-term outcomes (green).

The antecedent condition and the student profile or population reached were the least consistent components of the model. This is not necessarily problematic as one would expect differences in community or regional situations across the country as well as across the young people based in these regions. In fact, Project K is promoted as a community-driven programme; and, as Zhang’s (2011) interviewees also asserted, the programme content is meant to be modified to a certain extent in order to meet the unique needs of the different communities in which it is based. Nevertheless, the overarching theme that there is not enough support or opportunities for connectedness for young people at higher but not extreme risk seems to apply to the majority of these community contexts. Somervell’s (2011) research also suggested that the participant diversity may stem from an inconsistent application of the exclusion criteria across schools and across the regions. As a result, she made recommendations to improve the
standardisation of this process. If heeded, perhaps in time, alterations to this process will yield a more consistent profile of selected participants.

All the same, slight changes were made to the description of the population reached to incorporate the “theoretical” delimitations of the student profile. To be specific, the term *mid-risk* was included to denote the fact that while participants may have risk factors such as low self-efficacy, a lack of support, or a history of disruptive life events, they are not on the higher end of the risk continuum. Corroborating this idea is the notion put forward by some of Somervell’s interviewees that these participants are the silent sufferers who are troubled but are not necessarily noticed as such. This student profile description also aligns with the antecedent condition of the programme that support is lacking for youth “on the cusp of risk”.

The essential programme strategy (i.e. the Project K Experiential Learning Cycle) remains as it was because of the strong support several of these components had across all of the programme sources. It is evident that the different programme stages emphasise different types of experiences but it appears that they are all tied to one experiential learning cycle.

The influencing factors were also largely unchanged with the exception of one additional factor. *Cultural consideration* was introduced as a key programme influence after the results from the comparative analyses pointed to its significance. Every programme operates within its own particular culture and it is important to consider how this culture fits within the larger societal culture it is embedded in. This is particularly important in New Zealand’s bicultural society because, according to the Ministry of Youth Affairs (MYA)’s Youth Development Strategy Aotearoa, programme practices should align with the principles outlined in the Treaty of Waitangi (Ministry of Youth Affairs, 2002). FYD has reiterated the importance of this in the Policies and Procedures manual. The programme manuals also mention the importance of cultural sensitivity to participant backgrounds and cultural consideration is a component of the Project K’s mentor training programme (O’Neill, 2005). Furthermore, Qiao and McNaught (2007)’s
suggested that a participants’ ethnic background can influence programme outcomes and Hollis et al. (2011) also emphasized its importance, despite it not being a salient theme in the focus groups I conducted.

I chose to alter the format of the logic model outcomes not because there was a lack of consistency in the themes but because the expected timeframe of these outcomes was inconclusive; in fact, there was no specification of any timeframes for the outcomes in several of the data sources. Accordingly, I have collapsed the short-term and the intermediate outcomes (those that occur during the programme vs. those arising shortly after). For instance, motivation is now categorised with the other outcomes and resourcefulness is associated with recognising opportunities and resources. Some of the outcome labels were revised to better represent themes that were related but initially labelled differently. For example, the Project K staff and founders spoke about accountability and increased responsibility which was coded as an outcome of independence, whereas the youth participants also talked about gaining maturity; thus these were combined into one theme. Attitude change was acknowledged as a potential outcome by youth participants and it was a clear objective as outlined in the programme documents thus because the social conscience theme I identified in the focus groups was inherently about attitude change it was absorbed by this category. Health and fitness, although only having limited empirical support (Zhang, 2011), was also added because it was a topic mentioned in each comparative data source, and therefore an important part of the intended programme theory.

The desired long-term outcomes were not included in the preliminary model but because a positive future direction was a fairly prevalent theme across the focus groups and this desired outcome was corroborated by the document review, I have chosen to include it as an outcome arising as a result of sustaining the multiple preceding outcomes.

Ultimately, there was high concordance across the four data sources. This indicates that there is strong consensus with regards to how the programme is supposed to function. Interestingly, Zhang also noted consistency in the
programme conceptualisation across the different stakeholder groups she interviewed.

**Programme Implementation**

Consensus about how the programme is intended to be delivered across the organisation does not necessarily equate to congruence with actual programme delivery, a distinction described by Donaldson (2003) as the *conceptual vs. the action theory*. This is an important distinction to consider because, despite intentions, the practical constraints of the situation may create obstacles to implementing the programme as intended. To a degree this seems to be the case for Project K. For example, Somervell (2011) discovered that programme and school staff were restricted in the amount of time they could spend on the exclusion aspect of the Project K participant selection process. That, and the lack of knowledge the school personnel had about some of the eligible students, affected the way the exclusion criteria were applied. This then has implications for the profile of students who are selected for the programme. As a matter of fact, feedback on this component of the logic model by programme practitioners after a research dissemination workshop indicated that participants exhibiting high risk behaviours do get selected from time to time and this presents difficulties for facilitators.

The theory-building process unearthed other opportunities for programme improvement. Some Project K staff members spoke about the difficulty they had in engaging participants in the Community Challenge and this seemed to compromise the critical student-driven aspect of this component. Warren’s (2005) research indicated that this problem existed in previous programmes and thus is not a new concern. Additionally, Hollis et al. (2011) found that incidences of cultural insensitivity on the part of programme facilitators can create discomfort for some participants; these kinds of instances still arise on occasion despite the presence of some information about cultural awareness in the programme documents and mentor training programme. In a similar vein, the Facilitator Manual described potential problems that may arise when participants are reintegrated back into their classrooms and home lives following the Wilderness Adventure. It also describes several strategies for facilitating a smooth transition
but this issue was not raised in the staff and founder interviews even though this was a concern the youth interviewees discussed in Warren’s study. The stated objective that health and lifestyle behaviours (i.e. healthy eating, exercise, risk behaviours, and prosocial activity involvement) are expected to improve as a result of participation in Project K also appears to be somewhat tenuous.

Finally, it is important to reiterate that a number of participants did not complete the programme. This suggests that the programme experience was not satisfactory for all of the young participants. It is also possible that in providing an overall evaluation of their programme experience on the end of programme survey, Project K graduates may have overlooked any negative aspects of the programme. In fact, for some individuals, the programme may have had an inadvertent negative impact. An implementation evaluation of Project K would certainly need to account for these hidden perspectives and the potential for iatrogenic effects would need to be explored directly.

Those issues aside, there is also some evidence that the programme is being implemented as it should. The evidence from Qiao & McNaught’s and Warren’s studies shows positive changes in constructs related to self-concept, affect, connectedness, and goal-setting. Hollis et al.’s (2011) and Zhang’s (2011) interviews, though they were conducted with very small samples, revealed several of the same critical programme experiences, processes, and influencing factors identified by Project K staff and the founders. For example, at least one of the two studies spoke of the importance of challenge, experiential learning, teamwork, reflection, feedback, and the combination of autonomy support with scaffolding learning. The above demonstrates that a few aspects of the programme could be brought into closer congruence with the conceptual theory but what we do know about the implementation of the programme suggests it is being delivered, in large part, as intended.
The extensive literature reviews of youth development, adventure, service-learning, and mentoring programmes provided in Chapter Three provide the substance for the direct logic analysis, which, as explained earlier, is an evaluative process conducted to ascertain if a specific programme’s theory is consistent with relevant academic theories and empirical evidence (Brouselle & Champagne, 2011). To assess the surface legitimacy of Project K’s theory of change, each aspect of the Project K logic model was evaluated against the best practice guidelines, causal mechanisms, moderating influences, and outcomes discussed in academic literature on adventure programming, service learning, and mentoring (and presented in detail in Chapter Three). Table 6 provides a summary of the logic analysis findings. Each logic model component is included with a brief description of the associated themes; relevant support from the positive youth development, adventure, service learning, and/or mentoring programme literatures is provided alongside each component; relevant sources for the literature support are listed; and additional comments are made with regards to how this support compares with what is known about Project K.
<table>
<thead>
<tr>
<th>Logic Model Component &amp; Themes</th>
<th>Support from Literature</th>
<th>Relevant Sources</th>
<th>Additional Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANTECEDENT CONDITION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The social climate is not supportive of youth at higher than average but not extreme risk</td>
<td>The social climate of the 21st century in many industrialised nations is characterised by deterioration in connectedness, physical fitness and civic engagement; an alienated youth culture; lack of positive adult role models; and few opportunities for experiential learning.</td>
<td>Benson (1997) &lt;br&gt; Hopkins &amp; Putnam (1997) &lt;br&gt; Morrow &amp; Styles (1995) &lt;br&gt; Putnam (1996, 2001) &lt;br&gt; Rhodes (2005) &lt;br&gt; Roth &amp; Brooks-Gunn (2003a, 2003b) &lt;br&gt; Youniss et al. (2002)</td>
<td>The social conditions described in the literature fit with theme that support is lacking for young people generally in contemporary society</td>
</tr>
<tr>
<td><strong>STUDENT PROFILE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A diverse group of mid-risk 14-15 year olds who lack self-efficacy</td>
<td>The notion that all youth can benefit from asset-rich environments undergirds the positive youth development movement. Adventure and service learning programme results have had success with diverse youth populations, though mixed results do exist across age groups and ethnicity. Mentoring often targets at-risk youth but recent research illustrates that programmes are more effective for youth with some but not extreme levels of risk.</td>
<td>Billig (2000, 2004) &lt;br&gt; DuBois et al. (2011) &lt;br&gt; Hattie et al. (1997) &lt;br&gt; Jekielek et al. (2002) &lt;br&gt; Lerner, 2005 &lt;br&gt; Orren &amp; Werner, 2007 &lt;br&gt; Schwartz et al. (2010) &lt;br&gt; Wilson &amp; Lipsey (2000)</td>
<td>Project K appears to reach a diverse group of youth but targets the mid-risk sector of youth that generally gets overlooked by targeted services.</td>
</tr>
</tbody>
</table>
### Logic Model

<table>
<thead>
<tr>
<th>Component &amp; Themes</th>
<th>Support from Literature</th>
<th>Relevant Sources</th>
<th>Additional Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ESSENTIAL PROGRAMME STRATEGIES</strong></td>
<td>Experiential learning resulting from new and challenging real-world experiences, feedback, guided reflection &amp; transference</td>
<td>Conrad &amp; Hedin (1991)</td>
<td>The substantive aspects of the Project K experiential learning cycle are in close alignment with the cycle described in the adventure programming literature. The components of this cycle are reinforced in the community challenge and mentoring components. However, authenticity in the implementation of Community Challenge activities seems to be lacking in some programme sites.</td>
</tr>
<tr>
<td></td>
<td>Experiential learning (and the associated components of challenging and authentic experiences, psychological disequilibrium, feedback, reflection, and transference) is at the foundation of adventure programming. Service learning has an experiential learning focus with authentic activities and reflection described as particularly important. Authentic relationship-building activities and feedback are also integral to mentoring.</td>
<td>Dymond et al. (2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eyler (2002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hopkins &amp; Putnam (1997)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Luckner &amp; Nadler (1997)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Priest &amp; Gass (2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rhodes (2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rhodes &amp; Lowe (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seaman &amp; Gass (2004)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spencer (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spencer &amp; Liang (2007)</td>
<td></td>
</tr>
<tr>
<td><strong>TYPES OF ACTIVITIES</strong></td>
<td>Community exploration is one critical dimension of service-learning programmes and community action is the other. Physical challenges feature in all outdoor adventure programmes. Leadership activities and skill development are common characteristics of PYD programmes. Teamwork and building positive connections with others are essential</td>
<td>Terry &amp; Bohnenburger (2003, 2004)</td>
<td>The types of experiences Project K incorporates are all found in the PYD, adventure, service-learning &amp; mentoring programming literatures. The goal-setting process appears to have a stronger focus in Project K than in the youth programming literature.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Luckner &amp; Nadler (1997)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Garst et al. (2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Roth &amp; Brooks-Gun (2003a, 2003b)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Walsh &amp; Golins (1976)</td>
<td></td>
</tr>
</tbody>
</table>
## Project K’s Theory of Change

<table>
<thead>
<tr>
<th>Logic Model Component &amp; Themes</th>
<th>Support from Literature</th>
<th>Relevant Sources</th>
<th>Additional Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>aspects of all types of youth programmes.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### INFLUENCING FACTORS

**The Programme Environment: Intense, Safe and promotive of Autonomy in combination with Scaffolded Support**

The novel physical environment, challenging activities, and a supportive, small group setting are the defining characteristics of adventure programmes. Physical safety is a primary concern in adventure programming. More recently emotional/psychological safety has received attention in adventure programming. Physical and emotional safety is acknowledged to be essential to any PYD programme.

A youth-centred approach that encourages ownership and agency in decision-making but is balanced with guidance when needed is endorsed in the mentoring, service-learning, and PYD literatures.

- Davidson, 2004
- Davis-Berman & Berman (2002)
- Deane & Harré (under review)
- Dymond et al. (2008)
- Larson (2006)
- Lynch, 2006
- McKenzie (2000)
- Rhodes (2005)
- Rhodes et al. (2006)
- Rhodes & Lowe (2009)
- Terry & Bohnenberger (2003)
- Walsh & Golins (1976)

Project K appears to fare well in terms of producing the intense and safe environment; however, it seems that autonomy support and ownership could be more strongly emphasised in the Community Challenge in some regions.
<table>
<thead>
<tr>
<th>Logic Model Component &amp; Themes</th>
<th>Support from Literature</th>
<th>Relevant Sources</th>
<th>Additional Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual Characteristics: Facilitators, Mentors, and Participants (including Participant Engagement)</td>
<td>A substantial portion of the adventure programming and mentoring literatures deal with facilitator/mentor approaches and requisite skills and training. Participant characteristics including personal history, gender, age, ethnicity, and socioeconomic background are also touched on in each of the programming literatures but the empirical results of which facilitator, mentor and participant characteristics yield the most successful outcomes are mixed.</td>
<td>Billig (2004)</td>
<td>Like the patterns in the research literature, what is proposed about how individual characteristics interact with programme processes to generate different results is only discussed at the theoretical level. More empirical research is needed both within Project K and in the youth programming research.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Priest (1999b)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scale et al. (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sibthorp (2003a)</td>
<td></td>
</tr>
<tr>
<td>Logic Model Component &amp; Themes</td>
<td>Support from Literature</td>
<td>Relevant Sources</td>
<td>Additional Comments</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Interpersonal Dynamics: The Mentor-Mentee Match and the Group Dynamics</td>
<td>The quality of relationships with PYD staff and other adults can have a substantial influence on the success of a youth programme. In mentoring especially, the crux of a successful partnering is the bond between mentor and mentee. Cooperative learning tasks are central to adventure, service-learning and mentoring programmes. Support from team members in adventure programmes have been shown to influence individual and programme success.</td>
<td>Garst et al. (2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Martin &amp; Leberman (2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neill &amp; Dias (2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Parra et al. (2002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rhodes (2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rhodes et al. (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rhodes &amp; Lowe (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spencer (2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spencer &amp; Liang (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>As above, the interpersonal dynamics aspects discussed in Project K’s theory of change fit with the literature but more empirical research on these aspects is needed.</td>
</tr>
</tbody>
</table>
## Logic Model

### Component & Themes

<table>
<thead>
<tr>
<th>Ongoing Support: Multifaceted Continued Support (including External Networks) and Progression of Staged Programme Services</th>
</tr>
</thead>
</table>

### Support from Literature

- Programme dosage, both in terms of frequency of contact with programme services and the length of the programme, is an important moderating factor for all three types of youth programming and for PYD programmes in general.
- Long-term support is especially emphasised in the mentoring literature as iatrogenic effects have been evidenced with programmes shorter than 6 months.
- Family and community involvement are thought to enhance PYD programme quality and this has been shown to be true for youth mentoring. Genuine positive connections with community agencies are thought to be important to producing meaningful gains from service-learning programmes.

### Relevant Sources

- Catalano et al. (2004)
- DuBois & Neville (1997)
- Dymond et al. (2007)
- Hattie et al. (1997)
- Heinze et al. (2010)
- McBride et al. (2007)
- Roth & Brooks-Gunn (2003b)
- Scales et al. (2006)
- Scales & Roehlkepartain (2005)
- Urban (2008)

### Additional Comments

The disparate ideas of dosage and involvement of family and community in youth programmes are both related to the idea of ongoing support proposed in Project K’s theory of change. However, it is the integration of these aspects in combination with the progression of a young person through programme stages towards independence that appears to be a contribution unique to Project K.

Given the evidence that iatrogenic effects are associated with the early termination of mentoring relationships, it is concerning that some participants withdrew from the programme prior to graduation. The impact of the programme on these participants should be investigated.
### Logic Model

<table>
<thead>
<tr>
<th>Component &amp; Themes</th>
<th>Support from Literature</th>
<th>Relevant Sources</th>
<th>Additional Comments</th>
</tr>
</thead>
</table>

### OUTCOMES

**Positive Self-Concept; Knowledge & Skill Acquisition; Connectedness; Motivation & Achievement; Positive Outlook & Attitude Change; Recognising, Seeking & Utilising Resources & Opportunities; Maturity & Independence; Health & Fitness**<br><br>Meta-analyses of adventure and mentoring programmes and reviews of positive youth development and service-learning have evidenced consistent positive outcomes across self-concept, knowledge and skill development, connectedness, motivation, achievement, attitude change, and maturity outcome categories. The health outcomes associated with PYD, adventure and mentoring programmes are generally more focused on risk behaviours and psychological health.<br><br>Billig (2000, 2004) <br>Catalano et al. (2004) <br>Conrad & Hedin, 1991 <br>DuBois et al. (2002, 2011) <br>Eby et al. (2008) <br>Grossman & Tierney (1998) <br>Hattie et al. (1997) <br>Wilson & Lipsey (2000) <br>Yate & Youniss (1996) <br>Overall, we see the same pattern of outcomes result from Project K as those described in the literature with self-concept, knowledge and skill development, connectedness, motivation and achievement taking central stage over other types of outcomes, especially those associated with physical health.
Interesting Insights from the Literature

Several interesting insights arose from the direct logic analysis. A few of the programme development opportunities identified in the review of the Project K sources appeared to be matched by gaps or support in the research literature. For instance, cultural consideration appears not be a salient idea that informs the practice of Project K staff despite its inclusion in programme policy documents; however it also only receives brief attention at a theoretical level in the general youth programming and mentoring literature. In addition, it is largely ignored in the service learning and adventure programming literature.

Another opportunity for programme development identified during Project K’s logic model development was the position of the health and fitness outcomes in the theoretical model in that it is questionable as to whether these should be stated as primary outcomes. The findings of the logic analysis suggest that these types of outcomes also receive far less attention than others in all types of the positive youth programming literature. It appears that physical health outcomes may benefit from broader PYD approaches (Zhang, 2011) but generally these outcome are targeted by more specific interventions.

One important influencing factor that was highlighted in the service-learning literature but not in the Project K theory development process was the importance of having strong relationships with the community groups and members receiving services. Positive connections between the recipients of service and the young people coordinating and delivering the efforts are important aspects of creating authentic and meaningful service learning experiences (Dymond et al., 2007; National Leadership Council, 2004). This may be the thing that is missing from what appears to be a less than maximally impactful Community Challenge.

Grossman and Rhodes’ (2002) finding that mentees in the BBBS programme who received fewer than three months of mentoring were worse off than control participants also re-emphasises the importance of investigating the impact of Project K for those who did not complete the full programme. One Project K staff member mentioned the negative emotional impact that mentee abandonment by
mentors can have. The effects may be quite pervasive and long-lasting but our understanding of the impact programme non-completion has will remain elusive until an investigation is conducted.

On the other hand, it was encouraging to see that the idea that a gap exists in services for mid-risk youth in New Zealand, where Project K has found its niche, is now being picked up on in recent mentoring research. Mentoring programmes have recently shown that they can be most effective for this middle-risk group (DuBois et al., 2011; Schwartz et al., 2010) and DuBois et al. (2011) argues that mentoring alone is not appropriate for young people at extreme risk as these individuals generally need a different kind of intensive and therapeutic support. These concerns were noted early on in Project K’s development thus this reinforces that they have been on the right track for some time.

Finally, ongoing input as depicted in the Project K logic model extends the common notion of programme dosage. Rather than simply a matter of frequency of contact with programme services and programme length, ongoing input captures the idea that a greater impact can be made and transference enhanced when multiple sources of support are combined over a lengthy duration and services follow a progression that allows participants to transfer new lessons and skills within the programme setting and moves them towards independent learning.

Conclusions of the Logic Analysis

So does the Project K programme theory fit with the existing evidence base? The thorough model-building process used to establish Project K’s conceptual programme theory has revealed that the answer is a resounding ‘Yes’. Project K incorporates all critical aspects of adventure, service-learning, and mentoring programmes but does this as a unique staged progression of services that maximises interpersonal support and the dosage of programme services to further increase the likelihood of sustained positive change. Areas for programme development do exist and these should be further researched as part of a programme implementation evaluation. It is also important to note that the
development opportunities are not unique to Project K; many also appear to be areas for research development.

**STUDY LIMITATIONS, CONTRIBUTIONS AND CONCLUSIONS**

Limitations & Future Directions

First, the analyses were conducted by only one coder (me) and thus the interpretations of the results are limited by my cultural lens, research influences, personal values and experiences (some of which are outlined in Chapter One). However, the interpretive nature of this study did not lend itself to involving multiple coders. Furthermore, the findings are strengthened through the triangulation of multiple data sources and because in qualitative research the researcher becomes the “measurement instrument” (Patton, 2002, p. 86), the use of one coder across the different data sources should yield greater consistency in the measurements. Programme stakeholders (i.e. FYD and Project K staff who were involved in the focus groups and some who were not) were also consulted about the final version of Project K’s theory and their feedback on the interpretations was solicited when the findings were disseminated during a recent oral presentation (as mentioned in Chapter Two).

Second, a common criticism of programme logic models is that they do not reflect the complex systemic nature of the reality (Donaldson, 2003). Donaldson (2003, 2007), however, asserts that a more parsimonious version of the programme theory represented by the logic model is beneficial in that it helps to focus individuals on what are thought to be the core elements and as such it facilitates communication about the programme. I have endeavoured to reflect the complexity of the programme theory by including a cyclical process and depicting influencing factors as potentially intervening at any point in the process. I have also included some of the systemic influences, such as the family and community in the theoretical description that accompanies the model, and have avoided stating specific timeframes for the outcomes that are likely to arise as these may vary depending on the situation and the individuals involved. As a result, it is a very broad overview of the programme.
Related to the above comments, I initially attempted to qualitatively distinguish between mediating and moderating variables in the influencing factors component of the model. Mediating influences intervene between a predictor and an outcome such that the effect of the predictor on an outcome occurs either completely or partially via a direct relationship with the mediator, which has a direct relationship with the particular outcome (or a subsequent mediator). Moderators, in contrast, interact with a predictor such that the effect of the predictor on the outcome is conditional on the levels of a moderating variable (Baron & Kenny, 1986). This distinction is important as each has different implications for subsequent action but, as I found out, qualitatively determining the exact nature of the relationships between person, programme, and external factors is complex. The distinction is a statistical issue that is more accurately ascertained using appropriate statistical techniques and underlining this difference as part of Project K’s theory of change would have unnecessarily complicated the explanation for many stakeholders.

On another note, the Project K programme theory is heavily weighted towards the conceptual theory and, as such, it represents an idealised and descriptive version of how the programme operates. Although some insights into programme implementation were gained, a full understanding of the programme requires a more extensive implementation evaluation. An observational study of the programme would do much to contribute to this end as would testing the theoretical links statistically. The RCT study presented in Chapter Five contributes in part to the latter and, as mentioned previously (see Chapter Two in the chronology of the evaluation process), the logic model findings directed a follow-up quantitative study which focuses on measuring the links between programme experiences and new outcomes in addition to self-efficacy. The findings from this on-going study will also contribute to the statistical validation of the programme theory. This is further detailed below as I turn to the study’s contributions.
Study Contributions & Future Directions

Organisational Learning and Development

Taken as a whole, the contributions of this study far outweighed the limitations. In terms of organisational learning and programme development, I identified four main opportunities for programme development; these related to: an inconsistent student profile; cultural sensitivity and competence; engagement in the Community Challenge; and the tenuous position of physical activity as a programme objective. While, FYD has already taken some action on a few of these issues because they were raised in previous Project K studies (see Somervell, 2011 and Hollis et al., 2011 for details), the findings from this study substantiates what was found previously and the re-emphasis of these concerns has prompted further dialogue and action-oriented brainstorming sessions with a range of stakeholders (as indicated above and in the Chapter Two evaluation chronology).

The logic analysis process also pointed to an issue that, as mentioned above, has not been given much attention in Project K; that is the importance of building strong positive relationships with community members and organisations prior to or during the service learning experience. The literature suggests that activities that limit young participants abilities to cultivate connections with adults in the community (for example, doing a car wash for charity or picking up rubbish around the community) will be less authentic and meaningful (McBride et al., 2007) and, consequently, less effective in promoting their growth. What is more, in Chapter Three I indicated that there is the potential for adults to inadvertently further alienate young people from their communities through a service learning experience if the young people are not genuinely respected and engaged in the process (Fredericks et al., n.d.; Evans, 2007). Staff involved in these kinds of programmes (including Project K Community Challenge facilitators) need to be mindful of these issues and prepare community members carefully when bridging connections between young people and other adults in the community to ensure it is a meaningful and safe experiences. Inadvertent negative effects may also arise from some aspects of the Project K experience. This issue deserves more consideration both in research and practice.
These concerns have also raised some interesting evaluation questions that are either currently being explored or could be examined in future research. For example, does the inconsistency in the student profile across the programme regions really matter? Are there particular student characteristics in addition to those outlined in the exclusion criteria that impede programme success? Additional studies investigating how different individual characteristics influence programme processes and outcomes would help to clarify these questions, as the current literature offers few empirically-based conclusions. These questions are also linked to more specific questions about cultural differences. How does one’s cultural background influence programme success? Qiao and McNaught (2007) and Hollis et al.’s (2011) studies suggest they do but the small sample sizes associated with their results make this questionable. Hence, I examined the influence of gender and ethnicity with on the programme’s effect on self-efficacy outcomes with a much larger sample in Chapter Five.

Another question that merits greater attention in future research with Project K includes “What is the impact of the programme for those that withdraw prior to programme completion?” Although the staff and youth participants who were involved in this study were overwhelmingly positive about the impact of the programme, we are missing the voices of those who did not complete the programme. Youth mentoring research suggests that these individuals are most likely to be negatively impacted by their experiences; thus, it would be valuable to explore the potential for unintended negative effects more directly.

As mentioned earlier, the programme evaluation objectives could be altered to become more in line with the conceptual programme theory with regards to the health and lifestyle behaviours. Perhaps clarifying that these objectives are secondary to the primary outcomes relating to self-concept, connectedness, knowledge and skill development, achievement and motivation would help to close the gaps in the programme theory. This would also align with what has been found in the literature.
This process also contributed to the *evaluative direction* of Project K by underlining some critical programme variables that have not yet been extensively explored. Indeed, the findings from this study served as the impetus for a follow up study which has involved quantifying participant engagement and autonomy support across the programme stages and investigating how these constructs relate to other programme factors and outcomes. Moreover, as a result of this process, FYD’s Evaluation Coordinator and I decided to include outcome measures of connectedness (i.e. relatedness, Harré & Bullen, 2010) and sense of community (Peterson, Speer, & McMillan, 2008), and motivation (Guay, Mageau, & Vallerand, 2003) in the evaluation design for programmes operating from 2010 as we felt it was important to determine the programme’s effectiveness with regards to these other seemingly primary programme outcomes.

*Contributions to Cumulative Wisdom*

Often the findings from applied research studies are context specific in that they are utilised for decision-making with regards to a particular programme (Patton, 1997; Weiss, 1997b); however, this study has contributed in a number of ways to the *cumulative wisdom* on positive youth development and programme evaluation.

The unique way that Project K combines aspects of adventure programming, service learning, and mentoring has generated new ideas about how to maximise the positive outcomes resulting from youth development programmes. The idea of providing ongoing input via multiple sources of interpersonal support and a staged progression of programme services is an extension on the common notion of programme dosage that merits further exploration. A community service component was included in the original vision for Outward Bound and still features in some programmes (Hirsch, 1999). Its inclusion was meant to address Hahn’s dismay over the societal decline in compassion but as McKenzie argued in 2003, contemporary OB programmes seem to place much greater emphasis on individual growth to the detriment of the attention to service. In addition to engendering social conscience and compassion, Project K demonstrates that including a community exploration and service
component following an adventure stage can also serve as a means to progress participants from the novel environment to one that is more familiar while allowing participants to continue the transference of learning within a structured and supportive programme environment. Mentoring as a final adjunctive stage continues the progression towards independence and sustained outcomes. This is not to say that all programmes need to incorporate these particular strategies or even at this level of intensity. The environmental sustainability *Making A Difference* programme run by the Auckland City Council, for example, initiates their programme with a three day residential *hui* (gathering) in the bush for all selected youth participants. Then mentoring opportunities and volunteer activities are coordinated for those who are interested (Blythe & Harré, 2011); thus, the option to access ongoing input is still available.

The issue of ongoing input is closely linked to the idea of *participant reintegration* but as we learned with Project K, this requires more than continued support from personnel and progressive programme services; strategies must be implemented to ensure that participants do not feel abandoned following the programme. The mentoring literature touches on the importance of ending the relationship in an appropriate manner to avoid negative feelings on the part of the mentee (MENTOR, 2009) and DuBois et al. (2011) found that greater family and community involvement in youth mentoring can strengthen programme effects. This is also promoted in the general PYD programming literature (Durlak et al., 2007; Urban, 2008) and is of course the essence of service learning; however, we have seen that genuine community involvement can be difficult to coordinate. Moreover, the individuals who form a part of the young person’s micro-system external to the programme (e.g. family, peers, and schools) should be educated about the role they play in helping participants maintain their initial progress. This is especially so for programmes that remove participants from their normative environments (such as adventure programmes) as external networks can be very influential in easing the transition back into the normative environment (Deane & Harré, under revision).
Finally, this theory-building process has contributed to knowledge that is helpful in bridging the gap between evaluation theory and practice. This is a valuable contribution because, as stated in Chapter One, there have been multiple calls for more examples of how evaluation theory is enacted in the real world (Chen, 1994; Donaldson, 2007; Smith, 1993; Weiss, 1997b).

With regards to the programme logic modelling and theory-based evaluation literatures, the examples provided follow one of two general trends: programme activity-focused or programme process-focused models and the literature is replete with guidelines and examples of the former but not the latter (see The Health Communication Unit, 2001; Flex Monitoring Team, 2008; McLaughlin & Jordan, 1999; Porteous, Sheldrick, & Stewart, 2002; Fielden et al., 2007).

The choice may largely depend on the nature of the evaluand. For example, a programme with a very specific goal of improving reading abilities may benefit more from a theory that outlines the particular programme activities that promote reading skills rather than a theory that details the learning process involved. The broad aim of Project K, however, is to nurture the holistic growth of a young person over the course of 18 months. To this end, clarifying the process that cuts across all programme activities to achieve this overall aim is more helpful than if the model simply showed links between canoeing and the development of outdoor skills, for instance. Examples that illustrate how programme processes can function to support holistic psychological growth (rather than addressing particular component skills) are few and far between. Yet, it is knowledge about how to support these kinds of individual transformative processes that can really inform human service practice.

This study delineates a detailed description of the theory-building process employed to construct a programme process-focused logic model that does just that. This process incorporated various theoretical influences from the evaluation literature and was successfully enacted within the practical constraints of a real world evaluation context. It is an approach that achieved the dual aims of
producing a credible and evidence-based programme theory while giving a voice to multiple programme stakeholders (this is further explored in Chapter Six). As such, it is a valuable example of “thoughtful” theory-driven evaluation in practice.

Conclusion

As a final point, this evaluative endeavour to illustrate Project K’s theory of change drew on multiple data sources to generate a holistic picture of how Project K is presumed to generate positive changes in its youth participants. The process I employed sought to empower stakeholders by giving them a voice in creating the programme theory yet it also produced a highly credible product, as the consistency across the sources and the scientific evidence base suggest. This allowed us to determine the level of consensus across the organisation; it provided the organisation with a tool for future communication; it helped direct future evaluation activities; it identified opportunities for programme development; and it generated interesting insights about the youth development and programme evaluation literatures.
Chapter Five

EVALUATING THE BLACK BOX
PROJECT K’S IMPACT ON EFFICACY & ACADEMIC ACHIEVEMENT OUTCOMES

Self-belief does not necessarily ensure success, but self-disbelief assuredly spawns failure ~ Albert Bandura, 1997, p. 77

They are able because they think they are able ~ Virgil

THE CENTRALITY OF COMPETENCY-BELIEFS

The previous chapter presented an extensive programme theory-building process (see Figure 3) which can be used to produce a comprehensive, credible, and empowering evaluation. It went on to describe the “Exploratory Theory Development” stage and the first step, assessing consensus, of the “Theory Validation” stage. The programme theory produced from these efforts is offered as a tentative description of intended programme operations. However, further research is needed to confirm whether the proposed programme theory is playing out as it should in practice. The final (but on-going) phase of the “Theory Validation” stage, testing the hypothesised links, is required to ascertain the validity of the theoretical links. The information presented in this chapter contributes to this goal because it represents an investigation of the effectiveness of the programme on outcomes associated with self-concept (the most salient outcome identified through the logic model development process) using data collected as part of a randomised controlled trial.

Limitations notwithstanding, randomised controlled trials (RCT) are still believed to produce the most conclusive findings about cause and effect (and thus program effectiveness) because these designs provide the most accurate information about what would have happened to the individuals if there had been no programme (Cook, 2000; Lipsey & Cordray, 2000; Shadish, et al, 2002). Keeping
with the **credibility** aim of the *Thoughtful Evaluation* approach, programme effectiveness data generated from an RCT design coupled with hierarchical linear modelling (addressed further on) can yield rigorous findings. In essence, this is what is presented below, with a specific focus on self-efficacy and academic achievement outcomes. Because the links being tested are simply programme to outcome effects and no mediating paths are included, this may be considered a standard black box evaluation approach. That said, demographic moderators of the programme to outcome effects are also investigated; I will elaborate on this later when describing the details of study.

First, however, I will justify the focus on self-efficacy. Project K’s slogan “Maximising Youth Potential”, depicts a clear intention to produce flourishing young people who have the initiative to be all they can be. It is clearly framed as a competency and motivation-enhancing programme – not, for example, a “bond-building” or “esteem-enhancing” programme. Accordingly, Project K targets young people who have low self-efficacy with an aim of improving their beliefs of what they can achieve.

It is well-established that initiative and motivation are to a large degree contingent on a person’s **belief** in his or her capabilities, over and above actual skill and ability (Bandura, 1977; 1997; Eccles & Wigfield, 2002; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). Some theorists posit that **feeling confident and effective** in one’s actions (rather than actually **being** competent) is a prerequisite for well-being. In this way, **sense** of competency (or efficacy; see Harré, Tepavac, & Bullen, 2009) is seen to be an essential psychological nutrient at all stages of life (Ryan & Deci, 2000, 2002).

Furthermore, the significance of competency-related beliefs to well-being becomes even more apparent during adolescence (Bandura, 2006). As adolescents become more independent and goal-directed they require greater efficacy to self-regulate especially as academic pressures mount, and peer pressure increases (Bandura, 2006). Confidence to successfully navigate relationships is paramount as friendships increase along with opportunities to make romantic connections.
Competency beliefs related to career exploration and development also become more pertinent because young people must make decisions about what to do when they leave school (Bandura, 1986, 2006).

Unfortunately, many individuals experience a decline in their beliefs about their capabilities as they progress through school (Eccles & Wigfield, 1994; Schunk & Meece, 2006; Wigfield et al., 2002). This has been attributed to the inherent complexity and challenges of the adolescent environment (Caprara et al, 2008) including an increase in personal performance evaluations. Often these are referenced against those of classmates, especially in the competitive classrooms of middle and high school. Alternatively, it could simply be that as children mature they develop a more realistic understanding of the self. Naive optimism about one’s capabilities do tend to diminish as personal experiences in achievement contexts accumulate (Eccles & Wigfield, 1994; Wigfield et al., 2002).

Whatever the reason, competency-enhancing supports are essential to maintaining a healthy sense of personal efficacy during adolescence, especially when youth exhibit below average levels for their age group. Appropriately, efficacy is a theme that features strongly in the positive youth development (PYD) literature. As discussed in Chapter Three, PYD programmes strive to foster the 6 C’s, two of which are Competence and Confidence. Evidently, this entails nurturing young people’s self-beliefs especially with respect to what they can achieve. Aspects of efficacy are also apparent in the Search Institute ®’s Developmental Asset® framework which also informs PYD initiatives (Benson, 1997; Benson et al., 1998; Leffert et al., 1998). One of the four internal “building blocks to healthy development” is social competencies (Search Institute®, 2006). Assets within this category include cultural and interpersonal competence, as well as those for planning, decision-making, and self-regulation. Furthermore, in their review of high quality PYD programmes, Catalano et al. (2004) observed that efficacy is a commonly targeted outcome. Evidently, Project K aligns well with this aspect of PYD programmes.
To further demonstrate the importance of competency-related beliefs for human functioning and youth development and, more specifically, to provide a rationale for why self-efficacy is the focal outcome of Project K, I begin the next section with an overview several relevant social cognitive theories of motivation and achievement (beginning with self-efficacy theory) and illustrate why competency-related beliefs are fundamental to each. I then review the empirical evidence justifying the import of self-efficacy beliefs to three key domains of adolescent development. Next I discuss the promise adventure, service-learning and mentoring programmes hold for promoting adolescent efficacy beliefs and consider the way participant characteristics may interact with programme processes to generate different self-efficacy effects. Finally I conclude the introduction with the objectives of the randomised controlled trial study.

**EXPECTANCIES, VALUES AND HUMAN POTENTIAL**

Social Cognitive Theory (SCT) explains personal growth and development as the result of reciprocal influences between personal factors (cognitive, biological, and affective processes); behaviour; and the surrounding environment (Bandura, 1977, 1986, 1989, 1997, 2006). Relevant cognitions within the SCT framework include different types of expectancies and values. More specifically, a person’s expectations of success coupled with how much he or she values the task at hand (or the expected outcome) are central ideas in the motivational literature, and competency beliefs are a core part of this equation. Theories of self-efficacy, outcome expectations (Bandura, 1977, 1989, 1997, 2006), locus of control (Rotter, 1990), causal attribution (Weiner, 1985, 1986), expectancy-value (Eccles & Wigfield, 2002), and self-regulation (Zimmerman, 1989, Zimmerman & Cleary, 2006) all originate from the same overarching social cognitive framework and each draws on either expectancies, values, or both to explain why people pursue the things they do. Understanding how competency-related beliefs are linked to each of these theories substantiate their fundamental role in motivational behaviour.
Self-Efficacy Theory

First I describe self-efficacy theory as it is the core theory upon which Project K is based. Bandura describes self-efficacy beliefs as the judgements a person holds about his/her ability to enact the behaviours needed to successfully complete a task or attain a goal (Bandura, 1977, 1989, 1997, 2004, 2006). As such they are specific beliefs about tasks within particular domains (e.g. academic, social, athletic, and employment). At the same time, self-efficacy beliefs specific to one domain can generalise to other domains, especially when the required skills are similar (Bandura, 1977, 1997). Shell, Murphy & Bruning (1989) demonstrated this generality dimension when they found that self-efficacy beliefs for reading predicted writing achievement and vice versa.

Four main sources of self-efficacy have been identified in the literature. The more successful mastery performances we have, the stronger our self-efficacy beliefs for relevant future tasks whereas experiences of failure diminish confidence in our capabilities (Bandura, 1977, 1997; Schunk, 1989; Pajares, 2002; Zimmerman & Cleary, 2006). By observing others modelling the required actions of a task we can also vicariously build our confidence in executing it successfully ourselves (Bandura, 1977, 1997; Schunk, 1989; Pajares, 2002). The influence of vicarious mastery experiences is stronger when the model is similar to us (Bandura, 1997; Schunk, Hansen & Cox, 1987; Zimmerman & Kitsantas, 2002). To illustrate, Schunk, Hansen, & Cox (1987) found that in two different samples of primary school students who struggled with mathematics, observations of coping models who initially exhibited difficulties were judged to be more similar in competence and had a greater positive influence on the participants’ math-problem self-efficacy than observations of mastery models who solved problems with no struggles.

Verbal feedback about personal performance (Bouffard-Boucher, 1990) as well as feedback pertaining to the difficulty of the task (Maddux, Sherer, & Rogers, 1982) has also been shown to cause fluctuations in self-efficacy beliefs. Positive verbal feedback can engender stronger feelings of efficacy but it must be believable (Bandura 1977, 1997; Pajares, 2002).
Likewise, Bandura (1977, 1997) explains that when we become anxious just prior to a task, we often perceive the arousal as an indication that we should doubt our capabilities. His experiments with people who had a snake phobia have demonstrated that by managing the anxiety associated with upcoming feared tasks, we can increase our feelings of efficacy (Bandura, Adams, & Beyer, 1977; Bandura, Blanchard & Ritter, 1969; Bandura, Jeffery, & Gajdos, 1975). In consequence, Bandura asserts that anxious arousal is a fourth source (or threat) to self-efficacy (Bandura, 1977, 1997).

In a seminal article Bandura wrote in 1977, he argued that in terms of the cognitions that govern motivation and behaviour self-efficacy beliefs (or efficacy expectations) are the most pivotal. This is because individuals need to believe that they are capable of putting their knowledge and skills into practice to successfully achieve their personal goals. When they do, they are much more likely to commit to and pursue them. They are also more likely to set challenging goals, believe that positive outcomes will result from their actions, perceive obstacles as surmountable, and thus persevere when they encounter such barriers (Bandura, 1977, 1989, 1997, 2004, 2006).

Outcome Expectations, Locus of Control and Causal Attribution

Rather than focusing solely on expectations about how one will perform on an upcoming task (efficacy expectations), some theorists focus on expectations about outcomes. Outcome expectations are judgements about the consequences (whether advantageous or aversive) resulting from particular behaviours (Bandura 1997). For example, if Johnny does well on a test, he expects a good mark and praise from his mother. Both are positive outcome expectations that provide an incentive for Johnny to study hard. Yet, this does not reflect whether or not Johnny believes he can perform in a manner that will generate these outcomes. The two are related, however, as stated above.

Rotter’s (1990) concept of locus of control similarly concerns the outcomes of behaviours. He noticed that some people have a generalised tendency to believe outcomes are controlled by external factors like luck. Others believe things like
ability and effort (internal factors) play a large role in personal success. A Locus of Control approach assumes that these generalised expectations arise from previous reinforcement histories (Rotter, 1990) and those with a strong internal locus of control are more likely to be self-directed than those who tend believe outcomes are the result of external factors (Zimmerman & Cleary, 2006).

Weiner (1985, 1986) followed a similar train of thought when exploring why people naturally seek explanations for the outcomes they experience. He argued that this allows individuals to make adjustments to strategies following negative outcomes or to understand and retain those that are successful. Common attributions within achievement contexts include ability, effort, task difficulty, and luck (Weiner, 1985; 1986). Weiner recognised that these causal attributions do vary in terms of internal and external controllability but that they also vary on other dimensions. He, therefore, incorporated aspects of Rotter’s locus of control theory but argued that the locus of the attribution (internal or external) could be separated from a controllability dimension and they can vary in stability. Later he described other potential dimensions: globality (whether the cause is consistent over time and how much it generalises to other situations) and intentionality (whether the cause occurs out of volition). According to Weiner (1985) these attributions generate different emotional responses which are linked to value judgments and together they direct motivation.

Expectancy-Value Theory

Like Weiner, the “Eccles et al.” model of expectancy-value theory emphasises the dual role that expectancy beliefs and values have in driving behaviour (Eccles & Wigfield, 2002; Wigfield, 1994). For them, expectancy beliefs consist of expectations of success based on anticipated performances within a particular domain and the perceived difficulty of the task whereas “value” refers to judgments of the envisioned task. Judgments include the personal importance of success; the likely utility for immediate or future goals; intrinsic enjoyment or interest; and the cost associated with the negative emotions the task may generate, the amount of effort expended and the lost opportunity of choosing one option over another.
Self-Regulation Theory

According to Zimmerman and his colleagues the likelihood of success is greatly enhanced when individuals are self-regulated (Zimmerman, 1989; Zimmerman & Cleary, 2006). Self-regulation is a cycle of self-monitoring, self-evaluation, self-reflection and self-reaction. It aids decisions about which goals to pursue and which strategies to employ, adapt or forego prior to and during goal pursuit.

Many aspects of the above theories are incorporated in the description of their self-regulation cycle (Zimmerman & Cleary, 2006). Initially, individuals reflect on how much they value the task at hand (task value), what positive or negative outcomes may result from their attempt (outcome expectations), how much control they have over the outcome (personal control), how difficult the task is (task expectations), and how well they believe they can perform (efficacy expectations) before deciding whether to pursue a goal. Then, good self-regulators will employ particular strategies (e.g. organising information, careful planning, time-monitoring, setting proximal goals, memory rehearsal, etc.) and monitor their progress against goal standards. Depending on how they judge they are doing their strategies may change (i.e. if individuals believe they are meeting or exceeding the necessary requirements they are likely to persist with the task whereas when people feel they are underperforming they are more likely to give up). Once the final outcome is determined, people will make a causal attribution and will react to this in a pleasant or negative manner. These reactions, attributions, and the final performance assessment feed back into decisions about similar tasks when people encounter opportunities in the future (Zimmerman & Cleary, 2006).

The degree to which one strives for achievement and success depends on multiple socio-cognitive factors and the above theories demonstrate that competency-beliefs connect them all. Furthermore, a large evidence base supports this. Self-efficacy measures have been created and validated (both for their psychometric properties and for their ability to predict important well-being outcomes – see Pajares & Urdan, 2006) across many domains of human
functioning. To illustrate, increases in self-efficacy have been associated with positive changes in health behaviour (Holden 1992; Schwarzer & Luszczynska, 2006), academic achievement (Bandura, Barbaranelli, Caprara & Pastorelli, 1996; Carroll et al., 2009; Phillips & Gully), athletic performance (Feltz & Magyar, 2006), career identity development (Rogers & Creed, 2010; Stringer & Kerpelman, 2010; Taylor & Betz, 1983; Taylor & Popma, 1990), better psychological adjustment (Bandura et al., 1996; Connolly, 1989; Caprara, Gerbino, Paciello, Di Giunta & Pastorelli, 2010; Kim & Cicchetti, 2003) and fewer maladaptive social behaviours (Allen et al, 1990; Caprara et al., 2010), to name a few.

To conclude, this shows that Project K’s focus on self-efficacy is well-justified by psychological research and, as discussed previously, self-efficacy is generally thought of as a context-related concept. Accordingly, I next draw attention to what is known about self-efficacy within academic, social, and career development contexts because, as alluded to in the introduction, thriving in early adulthood necessitates competency-building within these three domains.

SELF-EFFICACY WITHIN ACADEMIC, SOCIAL AND CAREER DEVELOPMENT SPHERES

The Academic Context

Self-efficacy has been studied in many academic contexts including grade attainment (Mone, Jeffries & Baker, 1995), math (Pajares & Miller, 1994), science (Britner & Pajares, 2006), reading, and writing (Shell, Murphy & Bruning, 1989). The results of these studies generally point to the importance of efficacy beliefs for scholastic achievement. Multon, Brown, and Lent’s (1991) meta-analysis of 38 samples demonstrated that self-efficacy perceptions account for approximately 14% of the variance in academic performance. Studies conducted since then continue to substantiate this link (Bandura et al., 1996; Caroll et al., 2009; Phillips & Gully, 1997).

The study of academic self-efficacy can be divided into two main parts. The studies mentioned above generally refer to efficacy beliefs for particular subject areas (e.g. math knowledge and skills) but several theorists suggest that abilities
such as focusing on schoolwork, avoiding distractions, and completing homework on time are also incredibly important (Caprara et al., 2008; Pajares, 2002; Pastorelli et al., 2001) as these skills cut across all content areas. This subtype of academic self-efficacy has become known as self-regulation for learning efficacy. In consequence, Bandura and his colleagues in Italy have utilised a measure of academic self-efficacy that combines beliefs associated with various academic subjects with those for self-regulated learning on several different occasions (Bandura et al., 1996; Bandura, Pastorelli, Barbaranelli & Caprara, 1999; Carroll et al., 2009). Not surprisingly, higher scores on this measure have been linked to higher scholastic achievement (Bandura et al., 1996, 1999, Carroll et al., 2009) and greater educational aspirations (Bandura et al, 1996; Carroll et al., 2009). The self-regulation for learning subscale alone also predicted achievement and lower school dropout longitudinally using measures collected at the end of junior high and high school in a sample of 13 year olds from Italy (Caprara et al., 2008).

The Social Context

Measures of social self-efficacy reflect an individual’s belief in their capability to communicate with others, build relationships, manage interpersonal conflict, and assert personal viewpoints. It has consistently been linked to better psychological and behavioural adjustment, for example lower levels of depression (Bandura et al., 1996; Caprara et al., 2010; Hermann & Betz, 2006), loneliness (Hermann & Betz, 2006), withdrawal (Connolly, 1989; Wichmann, Coplan, & Daniels, 2004), other internalising symptoms (Connolly, 1989; Kim & Chichetti, 2003) and delinquent behaviour (Allen et al., 1990; Caprara et al., 2010).

An interesting study by Desivilya and Eizen (2005) showed how social self-efficacy may engender a more collaborative and productive interpersonal style. They examined 13 groups of young adults participating in a year-long voluntary community service residential Scout management programme in Israel. They were observed on engagement vs. avoidance and constructive vs. destructive dimensions of conflict resolution when working with their teams. Individuals with high social self-efficacy were more likely to employ integrative strategies that were
both engaging and constructive in comparison to dominating, avoidant, or overly compromising styles.

The Relationship between Academic and Social Self-Efficacy Beliefs

According to Patrick (1997) one’s sense of social competence is also intimately connected to success in academic contexts. Her review illustrates that individuals who are more socially integrated, demonstrate social responsibility, and proper social conduct have better school adjustment and thus achieve more and are less likely to drop out. She proposes that the links between academic and social competence emanate from similarities in self-regulation across both domains. Indeed, a study by Vecchio et al. (2007) revealed that social self-efficacy contributed to higher academic achievement independently of academic self-efficacy. Bandura et al (1996) also discovered that while social self-efficacy was not directly related to academic achievement, it was indirectly associated via effects on academic aspirations and depression. Academic self-efficacy has also been associated with the same or similar outcomes as those predicted by social self-efficacy. For example, lower levels of depression, greater prosocial (Bandura et al., 1996; 1999) and less problem behaviour (Bandura et al., 1999; Carroll et al., 2009).

At the same time, some aspects of social self-efficacy can be maladaptive for academic outcomes. When academic and social self-efficacy were investigated together, Carroll et al. (2009) found that social self-efficacy was negatively related to academic achievement in a group of high school Australian youth. This may be because young people associate high academic achievement with lack of popularity at times. It may also be that individuals who are highly committed to academic endeavours have less time to participate in social activities and thus have less opportunity to develop social skills (Carroll et al., 2009) and vice versa. It is true that several other studies described in this review have not found this association, yet Carroll et al.’s findings indicate that further investigation around the interrelationship of academic and social self-efficacy is warranted.
The Career Development Context

According to Erikson (1968), developing a vocational identity is one of the central tasks of adolescence. The solidification of personal strengths and interests during this life stage informs career preferences and beliefs about what is realistic. In one study of note, Bandura et al. (2001) investigated the links between the academic and social efficacy beliefs and self-efficacy for a variety of tasks found across six different occupational themes. Italian adolescents with high academic self-efficacy were confident they could function within the majority of occupational categories (i.e. science and technology, education and medicine, artistic and literary fields, and business and social services). Additionally, social self-efficacy was indirectly linked to confidence in several occupations through its association with personal aspirations. A year later, the researchers found that the young people’s career choices aligned with their efficacy beliefs for similar occupations measured the year before.

Lent, Brown, and Hackett have done extensive work on understanding how career interests develop and how career paths are then forged. They proposed that career identity development emerges from the interaction of personal factors (such as gender, ethnicity and socioeconomic status), social cognitive factors (efficacy, outcome expectations, and goals), and surrounding contextual factors (barriers and supports). They termed their framework Social Cognitive Career Theory (SCCT) (Lent, Brown, & Hackett, 1994).

Since proposing SCCT, numerous studies have investigated the links between efficacy beliefs and vocational identity. In consequence, we know that higher levels of career decision-making self-efficacy are linked to lower levels of indecision about university course specialisations (Taylor & Popma, 1990) and careers (Taylor & Betz, 1983; Taylor & Popma, 1990), greater career exploration (Rogers & Creed, 2010; Stringer & Kerpelmen, 2010) and intentions to explore career options (Ochs & Roessler, 2004), career planning (Rogers & Creed, 2010), and career identity commitment (Stringer & Kerpelmen, 2010).
In sum, efficacy beliefs have been consistently linked to healthy functioning within three particularly relevant areas of youth development. Measuring the effects of Project K on efficacy for academic, social, career development tasks should therefore provide good indicators of whether the programme is functioning to promote participant well-being across these spheres.

**SELF-EFFICACY ENHANCING INTERVENTIONS**

Several interventions have targeted self-efficacy beliefs. In fact, Bandura’s first studies of self-efficacy focused on the efficacy beliefs of people with a snake phobia within the context of clinical intervention (Bandura, Adams, & Beyer, 1977; Bandura, Blanchard & Ritter, 1969; Bandura, Jeffery, & Gajdos, 1975). Bandura demonstrated over multiple experiments that personal experiences of mastery, observations of similar models coping with increasingly threatening stimuli, and learning strategies to deal with anxious arousal (as discussed previously) led to increases in participant coping self-efficacy when dealing with actual snakes. In one study this effect also generalised to increased coping efficacy in anxiety-provoking public speaking situations (Bandura, Jeffrey, & Gajdos, 1975).

In 2008, Hyde et al. systematically reviewed 10 interventions to improve self-efficacy for dealing with addictive behaviour and the overall pattern demonstrated that self-efficacy within the context of addiction can be improved. Results from other studies also suggest that self-efficacy can be increased by intervention programmes (see Jerusalem & Hessling, 2009; Smith, 1989). Of particular interest here is whether adventure programming, service-learning and mentoring have been successful in improving self-efficacy.

**Adventure, Service-Learning, and Mentoring Programmes**

As stated in the introduction, PYD programmes should be promising avenues for increasing self-efficacy. Adventure, service-learning, and mentoring programmes (the core components of Project K) seem particularly well-suited for generating these increases. Ewert (1989), for instance, explained that self-efficacy is one of the central theories used to inform adventure programming initiatives. Adventure programmes provide participants with experiences of personal mastery.
in a challenging physical and social environment (Ewert, 1989; Deane & Harre, under revision). This is coupled with explicit feedback from instructors and peers (Deane & Harre, under revision) and the facilitation of what is essentially a self-regulation cycle. Hattie et al.’s (1997) meta-analysis of 96 adventure programme outcome evaluations confirms that these initiatives can have positive effects on self-efficacy.

Service-learning too provides challenging activities in interpersonal, decision-making, and project planning situations. In high quality service-learning programmes, students are encouraged to voice their opinions and to own decisions that can have a real impact on their communities (Dymond et al., 2008; Fredericks, n.d.; National Youth Leadership Council, 2004). Duffy and Raque-Bogdan (2010) suggest that service-learning programmes can be especially helpful in fostering career decision self-efficacy and results from their studies show that having a higher service motivation is linked to higher efficacy for make decisions about one’s future career. Conrad & Hedin (1991) summarised the effectiveness evidence for service-learning programmes from quasi-experimental studies and found studies reporting positive effects for social competence, feelings of mastery, interpersonal and political efficacy.

Models of mentoring also illustrate that having a mentor can lead to competency building across multiple domains (Rhodes, 2005; Rhodes et al., 2006; Rhodes & Lowe, 2009; Rhodes & DuBois, 2008). This likely occurs through joint activity which pushes a young person past their previous levels of achievement leading to additional experiences of mastery and a positive role model who provides vicarious learning opportunities and explicit verbal feedback. The two meta-analyses of mentoring programme evaluations led by David DuBois (DuBois Holloway et al., 2002; DuBois, Portillo et al., 2011) mentioned earlier also demonstrated that some mentoring programmes do improve competency-related outcomes, at least to a small degree.

Hence, there is some evidence indicating each of these programme types have been effective in increasing efficacy and/or broader competency beliefs.
However, the evidence for long-term effects remains thin, as does that examining important moderating effects like gender, ethnicity, and SES status. I turn to a discussion of these next.

**GROUP DIFFERENCES IN EFFICACY BELIEFS**

In Chapter Three I indicated that participants enter youth programmes with varied backgrounds and characteristics, and differences in gender, ethnicity, and socio-economic location may influence how they respond to programmes. I concluded that the empirical evidence is inconclusive even in terms of general outcomes. Nevertheless, here I consider the theoretical reasons why these demographic characteristics may yield differences in self-efficacy outcomes.

That self-efficacy beliefs are influenced by our social environment is evident from the role of feedback from others. This feedback may come in the form of the expectations about our abilities. The aspirations and expectations teachers and parents have of young people have in fact been shown to influence the efficacy beliefs and consequentially the performances of young people (Bandura et al., 2001; Schunk & Meece, 2006).

Unfortunately, often the expectations others convey to young people (whether intentional or not) endorse stereotyped beliefs about the subgroups to which they belong. Interestingly, negative stereotype beliefs have been proposed as an explanation for gender (Pajares, 2002), ethnic, and even SES differences in competency-related beliefs (see Wigfield et al., 2006) as well as differential responsiveness to cross-race mentors (Liang & Grossman, 2007; Sánchez & Colón, 2005).

Further, gender and ethnic differences in career decision self-efficacy and career choice trajectories seem to support this proposition (Bandura et al., 2001; Betz & Hackett, 1981; Gloria & Hird, 1999; Tang, Fouad & Smith, 1999). For instance, in samples of both American college students (Betz & Hackett, 1981) and Italian adolescents (Bandura et al., 2001), females reported higher efficacy for occupations that are traditionally occupied by females and less efficacy for those
traditionally associated with males. In terms of ethnic differences, Gloria and Hird (1999) proposed that the anxiety associated with racial and ethnic prejudice may influence the efficacy beliefs of racial and ethnic minorities. Their results showed that in comparison to White undergraduate students, those identifying with a racial or ethnic minority had lower career decision-making self-efficacy.

Brown et al. (1999) demonstrated that it is likely the interaction of SES and ethnicity that influences career decision making self-efficacy. When comparing career decision self-efficacy across urban and suburban high school students (urban contexts being associated with greater poverty and higher numbers of minority students), they discovered that career decision self-efficacy was linked to the distribution of one’s ethnicity within the educational setting. Minority students who formed the majority population in schools within urban environments had higher self-efficacy than Caucasian students in urban schools but Caucasian students from suburban schools had higher self-efficacy than minority students who formed a smaller proportion of these school settings.

Wigfield et al. (2006) suggest that along with conveying stereotype consistent expectations, young people are likely not encouraged to pursue activities that are associated with a negative stereotype of their gender or ethnic identity (e.g. females may not be encouraged to try out for the football team or a young African American may be swayed from pursuing a math specialisation). As a result they have fewer experiences of personal mastery to inform their sense of efficacy in these areas. To compound the issue further, few role models of their gender or ethnicity exist in these fields thus vicarious experiences of mastery are also likely to be lacking (Wigfield et al., 2006).

These are concerning issues for programme delivery and they are certainly pertinent to Project K, as programmes are primarily co-ed and accommodate participants of different ethnic groups and socio-economic backgrounds. In Chapter Four we saw that Project K staff believe participant characteristics can substantially influence programme success, as can the facilitators and mentors. Furthermore, cultural consideration was recognised as important in the
programme documents but it was not explicitly mentioned by programme staff involved in the focus groups I conducted. This suggests they may not be aware of the importance of this in programme delivery. Facilitators and mentors should therefore carefully consider if they are encouraging development along stereotypic lines, especially in situations where the facilitators do not share the same gender, ethnic, or socio-economic background to the participants they are serving. It is therefore important to ascertain if these demographic variables do indeed produce differences the programme’s ability to improve self-efficacy.

**THE CURRENT STUDY**

Thus far, this chapter has established that a strong sense of competency is fundamental to thriving in adolescence; that self-efficacy beliefs, in particular, are pivotal determinants of motivation, action, and achievement and that they are fostered through experiences of personal mastery, observational modelling of mastery, verbal feedback, and the regulation of anxious arousal. Adventure, service-learning, and mentoring programmes generally incorporate these elements to enhance young people’s competency beliefs which generates further growth (Chapter Three provides detailed reviews of each programme type). Given that Project K incorporates these three types of youth development programmes and all of their essential elements (as described in Chapter Four) we should see improvements in efficacy beliefs as a result of programme participation.

Ascertaining whether or not Project K is in fact effective in this regard is essential for demonstrating FYD’s accountability, as self-efficacy is the key criterion for selection and consequently the focal outcome. Granted, Qiao and McNaught (2007) did evaluate the effect of Project K on self-efficacy and academic achievement outcomes but their findings were limited in several respects. First, they only analysed a sub-sample of programme sites (n = 8) thus the total participant sample was small (approximately 100 participants at one year post-programme). In contrast, for the current study, data for participants from 50 different programme sites were available.
Second, the analytic strategy they employed lacked rigor in that two sample t-tests were used to calculate differences within each group and the difference in differences method was used to compare the net gains between the Project K and Control groups. There are major limitations to this kind of approach (see Bertrand, Duflo & Mullainathan, 2004) as it removes much of the individual variance in participants’ scores and thus reduces the accuracy of the estimates produced. Moreover, Qiao and McNaught’s (2007) analyses did not account for the dependency in the observations (participants within one programme site may be more similar than those from other programme sites) which is a violation of the assumption of independence for general linear modelling techniques (more on this below). Accordingly, it is important to ascertain if their findings can be replicated with a wider population of eligible Project K participants.

In consequence, the main objective of this study was to establish if Project K is an effective means through which academic (or self-regulation for learning), social, and career decision-making self-efficacy can be improved using a larger sample of programmes and more rigorous analytic strategies than Qiao and McNaught employed. Because the links between academic self-efficacy and academic achievement are well-established and improving educational outcomes is another objective of the programme, I was also interested in investigating the effects of the programme on academic achievement outcomes. Specific predictions are presented following a description of the study measures and procedure.

A second objective of the study was to ascertain if Project K is differentially effective for different participant subgroups based on gender, ethnicity and socioeconomic location. New Zealand is becoming increasingly diverse, especially within the youth population (Dunphy et al., 2008; Ministry of Social Development, 2010) and Project K is a programme that accommodates any student that exhibits below average levels of academic and social self-efficacy, as long as this does not co-exist with other very high risk behaviours. Consequently, the majority of the programme groups are co-ed and encompass all of the main ethnic groups that create the diverse social landscape of New Zealand but participants within these subgroups may respond differently to the programme, as mentioned above.
Additionally, as was discussed in Chapter Three, the core principles of adventure, service-learning, and mentoring models of youth programming tend to be U.S. or British imports. Project K does encourage tailoring the programmatic content to the local community context in which it is operating; however, the overarching structure remains standardised.

What is more, a report produced by Te Röpū Rangahau Hauora a Eru Pōmare (2002) stresses the importance of investigating the effects and experiences of health and well-being outcomes for numerical minorities (particularly Māori) within research studies exploring these issues for broad New Zealand populations. Without looking into the potentially differential effects and experiences across ethnic groups, findings will inherently reflect trends for the dominant majority (Pākehā/New Zealand Europeans). This can indirectly lead to further oppression of ethnic groups who are in the numerical minorities as policies and programmes based on this evidence will likely fail to accommodate their unique needs. In fact, ethnicity effects are intrinsically linked to socioeconomic factors likely because of issues such as these. Because of the bicultural agreement delineated in the Te Tiriti o Waitangi/The Treaty of Waitangi, Te Röpū Rangahau Hauora a Eru Pōmare (2002) argues that it is a breach of the Treaty agreement to conduct New Zealand research without consideration of the effects and implications for Māori.

All of the above factors reinforce the need to understand whether or not Project K produces differential effects for participant subgroups based on gender, ethnicity, and the SES status of their communities.

**METHODS**

**Programme Selection**

From the beginning of the randomised controlled trial (RCT) in late 2004 until the completion of recruitment at the end of 2007, 84 Project K programmes\(^{15}\)

\(^{15}\) A Project K programme is identified by its associated school and year the programme was implemented. For schools where more than one programme is delivered within one calendar year, a letter is attached to the programme name and year to distinguish it from the other programmes implemented in that same school during the same year.
were delivered. The participants for these programmes were sourced from 42 different schools across 11 different regions of New Zealand. Programmes within one region were governed by a single Community Partner Trust and, generally, programmes started in April, July, or September of these years. Each Community Partner implemented between one and seventeen programmes (mean = 7.64 and median = 7) and of the 84 programme sites, 50 were selected for the RCT evaluation.

School involvement in the evaluation process was of course contingent on obtaining consent from schools principals. Other than this, no consistent strategy was employed for the selection of programmes for the RCT. Community Partners delivering a small number of programmes were required to include all of their programmes in the RCT. New programmes (i.e. sourced from a school for the first time) were also required to take part. Community Partners with large numbers of programmes were not required to evaluate all programmes due to concerns expressed by some programme staff about a heavy workload and the need to maintain good relations with the schools; thus, schools with longer histories of evaluation were usually not selected. A few programmes were dropped from the RCT after being selected because of extenuating circumstances. For example, programme personnel in one region were not able to recruit a sufficient number of control students from one school and in another instance a programme director did not follow the selection guidelines (Moore, 2011, personal communication). In saying this, it is important to realise that while participants within each programme were randomly assigned to either the programme or a control group, programmes were not randomly selected for the evaluation; thus, one should be cautious about making generalisations to all Project K programmes. That said, the bulk of programmes (60%) delivered during this timeframe were evaluated in this manner and the only obvious bias in selection is that longer-running programmes implemented by more established Community Partner Trusts are less likely to have been included.
Participant Selection

The participant selection process occurred in three stages. The first stage involved the distribution of Moore’s (2005) Self-Efficacy Questionnaire to all students in Year 10 (13-15 year olds). Selection was based on scores for the three subscales (academic, social, and help-seeking) of this questionnaire (to be described in further detail in the following section). Two teacher ratings for each student were obtained for three different statements. Each statement was thought to reflect one of the self-efficacy subscales: “Is self-motivated and consistently works close to ability levels” (academic); “Is able to communicate positively with teachers” (help-seeking); and “Interacts positively with peers” (social). Responses were indicated on a 6-point Likert scale with responses ranging from “Not well at all” to “Very well” for each question. The median was calculated for each of the nine scores and individuals with scores below the median were coded as a “1” and those scoring equal or above the median were coded as a “0”. The dummy-coded scores were then summed and students were ranked by their score out of nine. Based on this composite score, approximately 50 of the highest ranked students were considered for the next stage of selection.

Also noted earlier, resource constraints and limitations of staff expertise prevent the capacity of the programme to accommodate high-risk students. To restrict participation to students that Project K is equipped to support, a strategy to identify high risk students was necessary. For the second stage of the selection process, programme directors organised a meeting with a liaison team of two guidance counsellors at minimum, the Year 10 dean, and any other teachers who were familiar with the Year 10 students and were willing to participate. The team reviewed the list of potential participants and any students identified by team members as having a recent history of violence, substance abuse, suicide attempts, regular counselling sessions for serious problems, severe cognitive or learning difficulties, and/or who were recommended for exclusion by a guidance counsellor (based on knowledge of confidential information) were removed from the list of potential programme participants. Each parent and potential participant was also asked to confidentially indicate if these criteria were applicable during a
subsequent Project K information meeting. These individuals were provided with information about other support services if there was any indication that their needs were unmet (Project K, 2006).

At least 20 female and 20 male students from the remaining list and their families were invited to attend the information evening. If parental consent and student assent were obtained (See Appendix C for participant information sheets and consent/assent forms), individuals for whom the exclusion criteria did not apply (according to parent and student report) were entered into a participant draw and were randomly allocated either to the programme (n = 12) or a control group (n = 12) in the final stage of participant selection. Any remaining participants were retained as reserves.

**Measures**

**Student Demographics**

On each evaluation survey, participants were asked to indicate his or her gender (male or female) and one or more of the ethnic groups to which he or she belongs. The options for ethnicity included: NZ Māori, NZ European/Pākeha, Tokelauan, Other European, Fijian, South East Asian, Niuean, Indian, Tongan, Chinese, Cook Islands Māori, Samoan, Other Asian, Other Pacific Islands, and Other.

Ethnicity output was categorised in two ways in this study and both aligned with the operationalisation of ethnicity by Statistics New Zealand (Allen 2001; Statistics New Zealand, 2005), which recognises ethnicity as a self-identified construct and acknowledges the potential for the existence of multiple ethnic identities. In order to obtain enough statistical power to detect effects related to ethnicity, it was first necessary to compile the Pacific nations and Asian ethnic groups into omnibus Pacific and Asian categories, and the Other European Category was combined with the NZ European/Pākeha responses as is customary for the highest level of aggregation for coding ethnicity (i.e. Level 1) according to Statistics New Zealand (Statistics New Zealand, 2005). A participant’s ethnic identity was initially coded using the Single/Combination Response coding method which assigns an
individual to only one ethnic category but categories include options for ethnic combinations, for example Māori/Pacific (Allan, 2001). Unfortunately, coding participant ethnicity in this way greatly reduced numbers for the Māori, Pacific, and Asian categories and necessitated placing all ethnic combinations within the same multi-ethnic category. This consequently reduced the power to detect ethnicity effects.

As a result, ethnicity was also coded using the Total Response method whereby a participant was counted in each group he or she selected; thus the total ethnicity responses using this strategy are greater than the number of participants in the sample (Allan, 2001). A comparison of the descriptive statistics for ethnicity using both of the coding strategies provides some understanding of the degree of discrepancy between the methods and this should be considered when interpreting the results (See Table 9 in the Sample Characteristics section below). Analyses requiring inferential statistics were conducted using the output from the Total Response strategy as this maximized the power to detect effects.

School Decile Rating

The New Zealand Ministry of Education (MoE) allocates funding and resources to state and state integrated schools in New Zealand according to the average socio-economic conditions of a school’s catchment area. Schools provide MoE with a random selection of home addresses from their student population and decile ratings for the relevant meshblocks are calculated from five socio-economic indicators gathered in the most recent New Zealand Census. The applicable indicators are household income, occupation type, household crowding, parent educational qualifications, and receipt of income support. Decile ratings range from 1 to 10 with decile 1 reflecting the 10% of schools with the greatest proportion of students coming from low socio-economic areas and decile 10 includes the 10% of schools with a catchment area containing the lowest

---

16 State and state integrated schools are two of three types of schools in New Zealand. Both are funded by the New Zealand government whereas private schools receive only a limited amount of their funding from the government. State integrated schools were initially private schools that later became part of the state system (Immigration New Zealand, 2011).

17 A meshblock is a small bounded geographical location used for the New Zealand Census (Ministry of Education, 2011).
proportion of students from low socio-economic communities (Ministry of Education, 2011). A list of decile ratings of the schools included in the RCT for each relevant programme year was obtained from MoE.

*National Certificates of Educational Achievement Level 1 Credits and Achievement Status*

Introduced in 2002, the National Certificates of Educational Achievement (NCEA) is the main national qualification for secondary school students in the New Zealand. Students obtain NCEA credits for achieving academic knowledge and practical skill standards within each subject studied. Standards are assessed both internally and externally and students obtain credits towards three different levels of certification (NCEA Level 1, 2 and 3). Normally, students work towards the Level 1 certification in Year 11, Level 2 in Year 12, and Level 3 in Year 13. Certification depends on the number of credits obtained and the difficulty level of each standard met. At the time of data collection (i.e. 2004 to 2007), to achieve the NCEA Level 1 certificate students had to attain 80 Level 1 credits and of these eight had to be literacy credits and another eight needed to be numeracy credits. The Level 2 certificate required 60 credits at level 2 or 3 and 20 credits from any level and 60 credits at level 3 or higher\(^\text{18}\) plus 20 credits from level 2 or higher were needed for the Level 3 certificate (New Zealand Qualifications Authority, 2011).

As part of the school, parent and student consenting agreements (See Appendix C), permission was requested for Project K/FYD to obtain each participant’s (Project K and Control) NCEA results from the schools as they became available. As the majority of participants in the current sample had completed Year 11 of school but not Year 12 at the time the follow up data was being collected, the total number of NCEA Level 1 credits obtained as well as the achievement status for the NCEA Level 1 certificate (i.e. whether or not the certificate was awarded) were the two academic achievement outcomes selected for this study.

\(^{18}\) Under some circumstances, students are able to obtain credits for courses taken at tertiary institutions.
Self-Efficacy

The Self-Efficacy Questionnaire (Moore, 2005) used for participant selection also served as the primary baseline evaluation measure (at Time 1) and was distributed to Project K and Control participants again upon completion of the programme (at Time 2) and one year following programme completion (at Time 3). The academic and social self-efficacy items are closely aligned with the academic and social self-efficacy subscales of Muris’ (2001) SEQ-C and Bandura’s children’s self-efficacy scale (Bandura et al., 1999) but some items were adapted to suit the New Zealand youth context. Academic items primarily assessed the degree to which individuals feel they can successfully self-regulate their learning (e.g. “How well can you study when there are other interesting things to do?”) but also included an item about meeting teacher expectations (i.e. “How well can you succeed in satisfying your teachers with your schoolwork?”) whereas the social items pertain to relationship building, communication and assertiveness in classroom situations (e.g. “How well can you have a chat with an unfamiliar person your age?”). Both of these subscales contain eight items measured on a six point Likert scale (1 = Not well at all to 6 = Quite well). A third subscale measuring political and community self-efficacy was initially included in Moore’s questionnaire; however the items did not load cleanly on a single factor when analysed using a principal components analysis. Four items that did seem to support a third factor related to obtaining resources or assistance from adults (e.g. “How well can you get the information you need from adults?”); thus, it was re-named help-seeking self-efficacy. In her study, Moore (2005) obtained Cronbach’s alphas of .88, .85, and .82 for the academic, social and help-seeking subscales, respectively (see Appendix D for the full questionnaire).

Career-Decision Self-Efficacy. Unlike, the academic, social, and help-seeking self-efficacy subscales, participant responses regarding their career decision self-efficacy were only collected at the one year post-programme time point (Time 3). The decision to include this outcome measure for the one year follow up was made in 2007 following discussions within FYD about the value of preparing Project K
participants for the school to work transition and the need to measure constructs associated with this transition.

This measure was an adaptation of the original Taylor and Betz (1983) 50-item Career Decision Self-Efficacy Scale (CDSE). The original scale was formulated based on an integration of Bandura’s Self-Efficacy Theory and John Crites’ (1978) Career Maturity Model (as cited in Betz, Klein & Taylor, 1996; and Taylor & Betz, 1983). The measure functions to assess the confidence one has in executing tasks that facilitate decision-making for the selection of one’s future career in the five main competency areas of Crites’ Career Maturity Inventory: self-appraisal, occupational information, goal selection, planning and problem solving (Betz, Klein & Taylor, 1996; Taylor & Betz, 1983).

The scale measuring career decision self-efficacy for this study was adapted substantially from the original scale in order to fit the language and developmental level of Year 11/12 students in New Zealand. Taylor & Betz (1983) acknowledged that many of their scale items are geared towards tasks that are particular to young adults in college, university, or on a professional career trajectory and may not be appropriate for investigations of career decision self-efficacy in other contexts, for example vocational or technical training. As a result, they expected future users of the scale to make modifications as deemed appropriate for the specific context and objectives to be investigated. Participants in this study were asked to indicate how much confidence they have in executing eight different career decision making tasks. An example of an adapted item is “Identify employers and further education/training institutions relevant to your career path”. These items reflected all but the problem-solving dimension of Crites’ five theoretical competency areas. Excluding the problem-solving content of the original scale was appropriate as most of these items reflected difficulties encountered after having chosen a career path. These did not seem appropriate for the developmental stage of the participants in this sample.

Throughout the evaluation process, Project K staff voiced concerns about the length of the evaluation measures for the young people in their programmes
(Moore, personal communication) and seeing as shorter versions of the CDSE had been used without compromising the reliability or validity of the scale (Betz, Klein, and Taylor, 1996; Fouad, Smith, & Enochs, 1997), a shorter measure also seemed justified. For example, in addition to the widely used Betz, Klein, and Taylor (1996) 25-item short form of the CDSE (the CDSE-SF), a 12-item scale adapted for middle school students (Fouad, Smith, & Enochs, 1997) has also stood up to internal consistency and construct validity benchmarks. In the end, the career decision self-efficacy measure designed for the Project K evaluation was reduced to 8 items to match the length of the academic and social self-efficacy subscales. In the case of this subscale, however, response options ranged from “No confidence at all” to “Complete Confidence” on a 5-point rather than a 6-point Likert Scale. The original scale designed by Taylor and Betz (1983) provided response options on a 10-point scale; however the use of a 5-point response continuum became popular in the 1990s (Betz, Hammond, & Multon, 2005) and Betz, Hammond, and Multon (2005) demonstrated that the 5-point continuum was as effective in terms of internal reliability and criterion-related validity as the 10-point version.

Major scale modifications such as the ones just discussed and application to a new population and context necessitate a thorough assessment of the measure’s structure and internal reliability. Analyses verifying the component structure and internal reliability of the academic and social self-efficacy subscales are thus presented below.

Self-Efficacy Scale Content Validity and Internal Reliability

At baseline, 1092 participants completed Moore’s (2005) Self-Efficacy Questionnaire and these data were subjected to a principal component analysis with Varimax rotation to obtain further support for the validity of the three subscales. When extracting all eigenvalues over 1, the solution resulted in a three-component structure with all items except for one social self-efficacy item (i.e. “How well can you express your opinions when your classmates disagree with

---

19 Principal component analyses were employed to maintain consistency with the approach Moore (2005) used when assessing the psychometric properties of the original scale; however, principal axis factoring was used to cross-check the findings and results were very similar.
you?"
)
loading most strongly on the theoretically proposed components. An
examination of the scree plot, however, indicated that a two-component solution
was more suitable as a third component added only minimally to the variance of
sample items. The actual mean eigenvalues obtained were as follows: 1 = 7.55, 2 =
2.29, 3 = 1.10, 4 = .93, 5 = .78. Two components accounted for 49.15% of the
variance while three components accounted for 54.63%. The first component
contained all of the academic items; the second contained all of the social items,
although three of these had cross-loadings > .4 on the third component; and all of
the help-seeking items loaded most strongly on the third component, though each
of these also had loading > .3 on the first component (See Table 1 in Appendix E).
A parallel analysis was then conducted to verify support for either the two or three
component solution. A parallel analysis allows one to compare the eigenvalues
obtained from the actual sample to those generated for a random sample of
unrelated items with the same sample size and number of items as the actual
sample. As such, it is acknowledged to produce more reliable results than both an
examination of the scree plot and/or simply retaining all eigenvalues larger than
one (O’Connor, 2000). Results from the parallel analysis pointed to a two-
component structure (Mean eigenvalues: 1 = 1.25, 2 = 1.21, 3 = 1.17, 4 = 1.14, 5 =
1.11); the mean eigenvalues for the first two components of the observed sample
were above the mean eigenvalues and the upper limit of the eigenvalue confidence
intervals (i.e. the 99th percentile eigenvalues) for the randomly generated parallel
sample while the actual mean eigenvalue for the third component (1.10) dropped
below what one could expect by chance for three components (1.17).

Another principal components analysis was conducted; this time the
extraction of two components was specified. An examination of the item content
and Varimax-rotated component loadings revealed that all of the academic and
social items loaded strongly on two separate components and while the help-
seeking self-efficacy items loaded more strongly on the component with all of the
academic items, they also had loadings > .4 on the component with social self-
efficacy items (See Table 2 in Appendix E). Conceptually, it is easy to see why the
help-seeking items did not emerge as a discrete component but rather overlapped
Project K’s Effectiveness

across Component 1 which contained all of the academic items and Component 2 which included all of the social items. Two of the items relate to obtaining support in a school context and the other two items are inherently about communication with adults. For this reason and because the academic and social subscales have been well-validated in other studies (Bandura et al., 1999; Muris, 2001), all of the help-seeking self-efficacy items were removed from the subsequent analysis. The next solution also demonstrated two clear components: all academic self-efficacy items loaded strongly on the first component and all social self-efficacy items loaded strongly on the second (see Table 3 in Appendix E). Table 7 (below) presents the actual eigenvalues obtained from the latter analysis, the cumulative % of variance accounted for by these components, and the mean and 99\textsuperscript{th} percentile eigenvalues obtained from the parallel analysis confirming support for the two-component structure. The internal reliabilities of both the 8-item academic and 8-item social self-efficacy subscales were adequate with Cronbach’s alphas of .88 and .85 respectively; thus justifying the retention of the academic self-efficacy and social self-efficacy subscales.

Table 7. Comparisons of actual and parallel analysis eigenvalues for self-efficacy items at Time 1 (specifying a two factor extraction)

<table>
<thead>
<tr>
<th>Component</th>
<th>Observed Sample</th>
<th>Parallel Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Eigenvalues</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>1</td>
<td>6.05</td>
<td>37.81</td>
</tr>
<tr>
<td>2</td>
<td>2.28</td>
<td>52.07</td>
</tr>
<tr>
<td>3</td>
<td>.97</td>
<td>58.13</td>
</tr>
<tr>
<td>4</td>
<td>.77</td>
<td>62.92</td>
</tr>
<tr>
<td>5</td>
<td>.70</td>
<td>67.28</td>
</tr>
<tr>
<td>6</td>
<td>.68</td>
<td>71.51</td>
</tr>
</tbody>
</table>

To verify the component structure of the Self-Efficacy Questionnaire longitudinally, these same analyses were conducted with the self-efficacy data collected at the completion of the programme (n = 891) and one year following the completion of the programme (n = 473). Very similar results were obtained. Note that the adapted version of Taylor & Betz’ (1983) Career Decision Self-Efficacy scale was introduced to the outcome measures at this point; thus three components were supported, as expected (see Table 8 below and Table 4 in Appendix E). The first
component again consisted of all of the academic items; the second component was formulated from all of the career decision items; the third component was comprised of the eight social self-efficacy items.

Table 8. Comparisons of the actual and parallel analysis eigenvalues for self-efficacy Items at Time 3 (specifying a 3 factor extraction)

<table>
<thead>
<tr>
<th>Component</th>
<th>Observed Sample</th>
<th>Parallel Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Eigenvalues</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>1</td>
<td>9.31</td>
<td>38.81</td>
</tr>
<tr>
<td>2</td>
<td>2.48</td>
<td>49.13</td>
</tr>
<tr>
<td>3</td>
<td>2.30</td>
<td>58.72</td>
</tr>
<tr>
<td>4</td>
<td>.96</td>
<td>62.73</td>
</tr>
<tr>
<td>5</td>
<td>.88</td>
<td>66.37</td>
</tr>
<tr>
<td>6</td>
<td>.78</td>
<td>69.64</td>
</tr>
</tbody>
</table>

The internal reliabilities of the academic, social and career-decision self-efficacy subscales created from the three factors described above also justified the use of an average score for each of these subscales (Cronbach’s α = .90 for academic, .88 for social and .90 for career decision). Furthermore, as cumulative (voluntary) item non-response was less than 5% for any particular item at each time point, missing values were not imputed.

Procedure

The self-efficacy questionnaire distributed for the purpose of participant selection served as one of the baseline measures for evaluating program success. For the first stage of participant selection, the questionnaires were distributed to all Year 10 students during class time by a Project K programme director who was accredited by the FYD National Support Office to conduct the evaluation process or by a trained staff member nominated by the programme director. It was recommended that at least one teacher remained in the classroom during survey administration. After the final Project K and Control participants had been randomly allocated to their respective groups and within two months of completing the baselines Self-Efficacy Questionnaire Project K students began the induction phase of the programme and the Wilderness Adventure.
In order to maintain contact with Control participants over the course of the RCT and to acknowledge their contributions to the evaluation an active control group process was implemented. Control participants were invited on a one or half-day activity outing before Project K participants returned from the Wilderness Adventure. They were also invited to three meetings with the programme director over the duration of the programme. Food was provided at these meetings and these served as an opportunity to connect with the students, update their contact details, assess whether personal support was needed and if so to provide details of support resources. The post-programme measures were administered at the third meeting and post-programme measures for both groups were collected no later than two weeks prior to the Project K graduation ceremony.

To maximise retention rates for the one year post-programme follow up, Project K personnel were required to contact participants from both groups every three months to ensure that the correct contact details were available. Data collection for the one year post-programme follow up occurred between twelve and fourteen months after programme completion. Programme directors were responsible for organising and supervising the administration of these questionnaires at a school or possibly at a student’s home.

Although some programme directors did organise and execute the data collection for the one year post-programme follow up for both the Project K and Control participants, many had difficulties playing such an integral role in the evaluation process while simultaneously delivering new programmes; therefore, a decision was made for the FYD National Support Office to take over data collection at this time point for the majority of programmes in the RCT. The FYD National Support Office sought and obtained funding to provide participants with incentives; therefore, from this point in the data collection process onwards, each participant was offered a $20 voucher for a mobile phone credit or a gift certificate for a department store chain found throughout New Zealand. Participants were initially contacted by telephone to ensure contact details were correct and they were also asked at this time if they would mind taking part in the follow up survey. Upon agreement, the survey along with the $20 voucher was mailed out. A
follow up email was sent two weeks after the initial phone call if a questionnaire was not returned. A final follow up phone call was made one week following the email if FYD staff had still not received the questionnaire.

Year 11 NCEA Level 1 credits and achievement status for all participants were collected from the schools by either Project K or FYD National Support Office staff at the start of Year 12. Programme directors were responsible for sending all of the completed surveys and NCEA data to the FYD National Support Office after each data collection time point. I obtained this data from the FYD National Support Office after obtaining ethical approval to access the secondary database from the Human Participants Ethical Research Committee at the University of Auckland pending an agreement to the terms outlined in a Memorandum of Understanding by the FYD, the University of Auckland, and the research investigators (See Appendix A).

Although the RCT evaluation itself was not reviewed by an ethics board prior to implementation, FYD outlined strict procedures to ensure that the evaluation process did not violate any ethical standards. As discussed, informed consent and assent was obtained from all parties involved (school principals, teachers, parents, and students) prior to participation. All participants’ confidentiality was maintained – no identifiable information was indicated on any questionnaire; students were identified by a unique code and only a select number of FYD staff have access to the names matching these student codes and the associated data (See Appendix C for Project K participant information sheets and consent forms). Furthermore, all FYD participant information sheets and consent forms distributed to Project K stakeholders were reviewed by the University of Auckland Human Participants Ethics Committee prior to permitting my access to the secondary database.

**Study Predictions**

As stated earlier, the primary objective of the study was to ascertain whether or not Project K is effective in improving self-efficacy and academic achievement outcomes. In line with the rationale provided in the introduction the
following predictions are put forward (all statements refer to programme effects while statistically adjusting for gender, ethnicity, and school decile rating effects):

1) Project K participants will show an increase in their levels of academic and social self-efficacy from baseline to one year post-programme and the rate of increase will be greater than that of a randomised control group.

2) Project K participants will have higher levels of academic and social self-efficacy than the Control group immediately post-programme and this difference will be maintained one year-post-programme, adjusting for baseline levels of academic/social self-efficacy.

3) Project K participants will have higher levels of career-decision self-efficacy than the Control group one year-post-programme.

4) Project K participants will have a higher number of NCEA Level 1 Credits and a greater proportion will have achieved the NCEA Level 1 Certificate than the Control group one year-post-programme.

Again, a secondary study objective was to explore the interactions between programme effects and gender, ethnicity, or school decile rating; however, as the evidence for these moderators is mixed in relation to youth programming and differential effects may be contingent on variations in programme delivery factors, no specific predictions are put forward.

Sample Characteristics

Programme Level

The 50 Project K programmes included in the RCT sample were sourced from 38 different schools that varied across the entire decile rating continuum (range = 1-10). The mean decile rating for these schools was 5.82 (SD = 2.72) and the median was 6. Figure 6 shows the frequency distribution of decile ratings for all of the programmes included in the sample.
The baseline sample consisted of 516 Project K and 576 Control participants. As discussed, participants were in Year 10 of secondary school at programme start, and so were between the ages of 13 and 15. The Project K group contained 238 females and 278 males whereas there were 276 female and 298 male Control participants (two participants did not specify a gender).

Table 9 presents a comparison of the ethnic breakdown of participants within each condition according to the Single/Combination Output and the Total Response Output strategies. Note that using either coding strategy, participants identifying with a European ethnicity represented the great majority, followed by those identifying with a Māori, Pacific, and Asian ethnic identity. Recoding individual responses in the combined categories to counts for each separate major ethnic group resulted in much larger numbers in each of the major ethnic categories, as noted earlier. Approximately 20% of the sample reported identifying with more than one ethnicity for both the Project K and Control conditions. The most common ethnic combination for Project K and Control groups was Māori/European followed by Māori/Pacific, Pacific/European, Asian/European, and European/Other Ethnicity Not Identified.
Table 9. Comparisons of baseline frequencies and percentages for breakdown of ethnic categories across Project K and Control conditions using both the Single/Combination and Total Response output strategies

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Single/Combination Output</th>
<th>Total Response Output</th>
<th>$\chi^2$ for differences across Condition$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PK</td>
<td>Control</td>
<td>PK</td>
</tr>
<tr>
<td>European</td>
<td>312 (61.0%)</td>
<td>304 (52.8%)</td>
<td>395 (76.8%)</td>
</tr>
<tr>
<td>Māori</td>
<td>40 (7.8%)</td>
<td>63 (10.9%)</td>
<td>119 (23.2%)</td>
</tr>
<tr>
<td>Pacific</td>
<td>40 (7.8%)</td>
<td>60 (10.4%)</td>
<td>69 (13.4%)</td>
</tr>
<tr>
<td>Asian</td>
<td>13 (2.5%)</td>
<td>19 (3.3%)</td>
<td>23 (4.5%)</td>
</tr>
<tr>
<td>Other Ethnicity Not Listed</td>
<td>8 (1.6%)</td>
<td>11 (1.9%)</td>
<td>10 (1.9%)</td>
</tr>
<tr>
<td>European/ Māori</td>
<td>60 (11.8%)</td>
<td>59 (10.2%)</td>
<td>-</td>
</tr>
<tr>
<td>Māori/Pacific</td>
<td>13 (2.5%)</td>
<td>16 (2.8%)</td>
<td>-</td>
</tr>
<tr>
<td>European/Pacific</td>
<td>11 (2.1%)</td>
<td>13 (2.3%)</td>
<td>-</td>
</tr>
<tr>
<td>European/Asian</td>
<td>4 (0.8%)</td>
<td>10 (1.7%)</td>
<td>-</td>
</tr>
<tr>
<td>European/Other Ethnicity</td>
<td>4 (0.8%)</td>
<td>9 (1.6%)</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>514</td>
<td>576</td>
<td>616</td>
</tr>
</tbody>
</table>

$^a$ A series of independent $\chi^2$ analyses were conducted using ethnicity data based on the Total Response method of coding.

**Preparatory Analyses**

To assess the degree to which the random allocation process functioned to create equivalent groups across the programme and control conditions at baseline as well as the effects of the non-random removal of Project K non-graduates (explained below) on the balance of participant characteristics across the two conditions, *chi-square* analyses were conducted to compare the final Project K and Control baseline samples according to gender and ethnicity, and independent samples *t-tests* were employed to calculate differences between conditions in Time 1 academic and social self-efficacy.

No significant differences were found between the Project K and Control groups with regards to the female to male ratio. There was a significantly greater proportion of participants identifying with a European ethnicity in the Project K group compared to the Control group ($\chi^2 (1, N = 1076) = 8.23, p. < .01$) indicating...
that more ethnic minorities belonged to the Control group than the Project K group at baseline (See Table 9 for Total Response Output percentages); however, the size of the effect was small ($\Phi = 0.09$). Results from the independent samples $t$-test analyses revealed that the 514 Project K and 576 Control participants did not differ statistically in mean pre-programme academic or social self-efficacy at programme start.

**Reasons for Incomplete Participant Data**

The number of individuals who participated in any part of this study was 1175; however, for several reasons not all of these participants were included in some or all of the study analyses including: Project K participants not graduating from or dropping out of the Project K programme; Project K and Control participants dropping out of the evaluation or programme directors not being able to contact evaluation participants; participants choosing not to respond to specific survey questions; and finally, school-specific survey items that were no longer applicable to school leavers at the final time point.

To clarify the final reason for incomplete data: Sample participants would have been between the ages of 15 and 17 at the time of the one year post-programme follow up. According to New Zealand law, students are only obliged to remain in school until the age of 16 and many graduate from secondary school at age 17. Accordingly, information pertaining to the current secondary school status of the sample participants was obtained during this data collection time period. Of the 596 students for whom this information was obtained, 107 had left secondary school. For these individuals, the academic subscale of the self-efficacy questionnaire was no longer relevant, thus it was removed from their follow up surveys. The smaller sample size for the academic self-efficacy score at Time 3 is reflective of this (see Table 6 in Appendix E).

Related to this, because the context of four of the social self-efficacy items were specific to an academic environment modifications to these items were required for the one year post-programme follow up measure in order to maintain relevance to school leaver participants. In consequence, four items which were
identical to the academic environment-oriented items with the exception of alterations to a few words, for example, exchanging “classmates” with “people your own age”, were added to the subscale. School students responded to 12 items for this subscale (the 4 new items plus the 8 original items) while school leavers responded to four of the original items plus the 4 new items (see one year post-programme survey for school students vs. school leavers in Appendix D). A comparison of school students’ social self-efficacy scores composed of the original 8 items to a score with the 4 original and 4 new items were made using a paired sample t-test. While there was a significant difference in these scores ($t (493) = -6.55, p < .01$) the paired sample correlation between these two scores was .97 suggesting high multicollinearity. In the end a decision was made to analyse the one year post-programme social self-efficacy outcome data using a score composed of the 4 original items and the 4 modified items for both school students and school leavers.

Table 10 provides a breakdown of the participant numbers associated with each type of non-inclusion and in the following section I discuss the degree to which these factors may influence the interpretation and generalisability of the results.
Table 10. Frequency and percentages for participant non-inclusion.

<table>
<thead>
<tr>
<th>Reason for Removal</th>
<th>Project K</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project K Participant Programme Incompletion</td>
<td>80 (13%)</td>
<td>n/a</td>
</tr>
<tr>
<td>RCT Participant Attrition at Time 2</td>
<td>45 (9%)</td>
<td>159 (28%)</td>
</tr>
<tr>
<td>RCT Participant Attrition at Time 3</td>
<td>207 (40%)</td>
<td>294 (51%)</td>
</tr>
<tr>
<td>School Leavers (Academic Self-Efficacy Subscale No Longer Relevant)</td>
<td>58 (17%)</td>
<td>49 (19%)</td>
</tr>
</tbody>
</table>

**Item/Subscale Non-Response**

<table>
<thead>
<tr>
<th></th>
<th>Project K</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0</td>
<td>2 (0.002%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0</td>
<td>2 (0.002%)</td>
</tr>
<tr>
<td>Career Decision Self-Efficacy</td>
<td>0</td>
<td>1 (0.001%)</td>
</tr>
<tr>
<td>NCEA Level 1 Credits*</td>
<td>126 (24%)</td>
<td>178 (31%)</td>
</tr>
<tr>
<td>NCEA Level 1 Status*</td>
<td>127 (25%)</td>
<td>178 (31%)</td>
</tr>
</tbody>
</table>

*These data were obtained directly from schools or the Ministry of Education.

An obvious corollary to the multiple reasons for non-inclusion was great variation in sample sizes across data collection time points as well as across variables. Tables 5 and Table 6 in Appendix E provide the frequency and percentage of Project K and Control participants associated with each outcome variable across participants in all of the programmes combined. The decline in participants over time also resulted in some variation in the sample characteristics across different variables and time points, as evidenced in Table 7 of Appendix E.

**Effects of Incomplete Data on the Interpretation of the Results**

To successfully graduate from Project K participants need to have accomplished at least three goals (one must be academic and another must be linked to physical health) and have participated in at least 75% of the activities. As a result, someone may participate intermittently until the end of the programme, but his/her status at post-programme will be registered as “incomplete”. Also, people may of course withdraw from programme at any time. These individuals were not included in the current analyses because they were exposed to different levels of exposure than those who participated more fully and this can lead to misinterpretations about programme effectiveness (Lipsey & Cordray, 2000). It is, however, helpful to know if these individuals differed systematically from those who had higher levels of exposure when making generalisations about the findings to the population of potential Project K participants.
To determine whether or not Project K graduates differed significantly from those who did not complete the programme, chi-square analyses were conducted to compare the gender and ethnic distribution\(^{20}\) in the sample of non-graduates to that of the final Project K participant sample. Independent samples t-tests were used to investigate differences between these two groups in their academic and social self-efficacy scores at baseline. The results of these analyses can be found in Appendix E.

As shown in Table 10, there was approximately a 9% attrition rate in the evaluation for the Project K group from pre to immediate post-programme measurements. In comparison, there was approximately a 28% attrition rate for the Control group over this duration. Attrition rates from pre-programme to the one year post-programme follow up, with regards to the self-efficacy data, were approximately 40% for the Project K group and 50% for the Control group\(^ {21}\). Chi-square and independent samples t-test analyses were again conducted to assess significant differences in participant attributes and baseline self-efficacy scores between those who were missing post-programme self-efficacy data and those with this data. The same analyses were conducted for missing self-efficacy data at the one year post-programme follow up time point. Results from these analyses can also be obtained from Appendix E.

**Caveats for Data Interpretations**

Taken together the results found in Appendix E and from the assessment of Project K and Control group equivalence at baseline indicate that a few caveats should be made with regards to the generalisability of the results described below when applied to all young secondary school students with below average academic and social self-efficacy. It appears that European students and those with higher social self-efficacy relative to other potential or actual Project K participants were more likely to remain in the programme, while Pacific Peoples were more

---

\(^{20}\) Ethnicity, in this case, was coded according to the Total Response Output method as described earlier.

\(^{21}\) This attrition was not additive to that described for the pre to immediate post-programme period as some participants who did not provide evaluation data at post-programme did respond to the survey at the one year post-programme follow up.
likely to drop out. Perhaps these results are indicative of the programme content not being a good cultural fit with the values of some Pacific students and it may be that it was more appealing to individuals with at least a moderate level of social confidence but these are purely speculations.

Furthermore, the Project K group had a higher proportion of European participants than the Control group at baseline and (for both the immediate post-programme and one year post-programme measurements) Europeans were more likely to respond to the self-efficacy survey, while Māori and Pacific participants were less likely to respond. Individuals with higher baseline social self-efficacy in comparison to their Project K or Control peers were also more likely to respond to the survey at the one year post-programme follow up. As a result, interpretations of the results as they pertain to the different ethnic groups and to individuals with social self-efficacy levels falling on the very lower end of the spectrum should be made with these considerations in mind.

**Analytic Strategy for Primary Results**

In comparison to the more widely used techniques following the rubric of General Linear Modelling (e.g. Ordinary Least Squares Regression and Analyses of Variance), large-scale multisite evaluations, such as Project K’s RCT, are better served by multilevel modelling strategies (in this case, hierarchical linear modelling or HLM) when analysing quantitative data because this class of statistical analyses offers greater methodological precision. HLM’s advantage emanates from the fact that multisite evaluations inherently produce a hierarchical data structure of participants nested within a programme site. This type of grouping structure accounts for the fact that participants clustered within one programme site are likely to be more similar to each other than they are to individuals from other sites. Similarities may stem from sharing the same school or community context and programmatic factors. These similarities introduce dependency in the data thus violating the assumption of GLM techniques that there is independence in the observations. Disregarding the grouping structure and analysing data at the lowest level can bias parameter estimates in serious ways.
leading to misinterpretations of the effects, such as the well-known ecological fallacy (Hox, 2002; Snijders & Bosker, 1999).

With HLM, the dependency is accounted for by simultaneously modelling variance at each level of analysis (e.g. the individual level vs. the group level). Additionally, by producing regression equations for each group, the reliability of the group parameters can be calculated. HLM weights these values accordingly when calculating the overall estimates; thus, larger groups with more consistent responses will have greater weight in the overall estimate calculations (Hox, 2002). Analysing data in this way also means that one can appropriately model the influence of Level-2 factors on Level-1 outcomes (e.g. programme characteristics such as instructor experience on participant outcomes such as individual self-efficacy) as well as cross-level interactions (Christ, Sibley, & Wagner, 2012; Snijders & Bosker, 1999; e.g. the moderating influence of school decile on the relationship between condition and career decision self-efficacy).

*Three Level Models*

Although individuals are a common unit of analysis at Level-1, repeated measures or observations may also be considered as Level-1 variables nested within individuals at Level-2 who may or may not be nested in programme sites at Level-3 (See Figure 7). The same logic as described above applies when analysing three-level models (Hox, 2002; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999); however, in the above example the variance would be partitioned into the programme site-level, the individual-level, and the measurement-level.

---

22 Of course, individuals nested within programmes which are nested within schools may also be analysed using a three-level model if there are enough cases at level-3.
Figure 7. Example of a three-level hierarchical data structure for a multisite evaluation where the full programme may fall under a national governance (like Project K) but programmes at each regional site are allocated to Level-3, individuals within each site are at Level-2, and repeated observations of/from these individuals are at Level-1.

An additional benefit of analysing repeated measures using HLM is that the participant attrition that often occurs at later measurement time points does not pose as much of a problem with HLM as it is with other statistical strategies (Field, 2009; Hox, 2002; Twisk, 2006). Participants with missing data at some time points will still be included in the analyses but estimations for these growth trajectories
will be weighted to account for the lower reliability in these estimates. In contrast, these individuals cannot be included at all when analysed using GLM techniques.

The Current Analyses

To provide a general overview of what was occurring between Project K and Control participants over time for the academic and social self-efficacy outcomes, I first report the growth trajectories of participants’ scores from pre (Time 1) to one year post-programme (Time 3) using three-level hierarchical linear models. The lowest level of analysis here being the repeated measures (i.e. Time), individuals were at the second-level, and programme sites\textsuperscript{23} at the third. Curvilinear effects in the trajectory of the outcome variables (i.e. academic and social self-efficacy) over the three time points were also investigated for the three-level models by including a quadratic term for the Time variable in the regression equation at Level-1. This demonstrates whether or not the linear trajectory accelerates or decelerates at any point, and the degree to which it does (Singer & Willet, 2003). Programme effects for both the linear and curvilinear trajectories were estimated while statistically adjusting for the main effects of gender and ethnicity (Māori, Pacific, and Asian)\textsuperscript{24} at Level-2, and school decile rating at Level-3. Outcomes were analysed separately (detailed model specifications to follow).

To obtain a more fine-grained understanding of the programme effects, two-level models with individuals at Level-1 and programme sites at Level-2 were then generated. As with the three-level models, outcomes were analysed separately. To determine immediate post-programme effects on academic and social self-efficacy, differences between Project K and Control participants at post-programme (i.e. Time 2 academic and social self-efficacy) were investigated, adjusting for the

\textsuperscript{23} Although different Project K programmes were delivered within the same location but in different years or at different times of the year I use the term programme site to refer to the 50 different Level-2 units to indicate that the control group belongs to the same site as programme participants.

\textsuperscript{24} The “European” and “Other” variables were not included as covariates in any analyses that incorporated ethnicity effects to avoid model multicollinearity (see Brambor, Clark & Golder, 2006). Furthermore, the main programme effects reflect trends that are biased towards the numerical majority and the secondary research question pertained to differential effects for the key numerical minority ethnic groups in New Zealand (see Te Röpū Rangahau Hauora a Eru Pōmare, 2002, for an argument supporting this rationale).
baseline measure of the relevant outcome (i.e. either Time1 academic or social self-efficacy). To ascertain if any differences were sustained one year after the programme, Project K and Control scores on these outcomes were compared at Time 3 adjusting only for the relevant baseline score. Then, to determine if there was an additional effect of the programme over the follow up period (i.e. between Time 2 and Time 3), Time 3 scores between the two groups were compared adjusting for both baseline and post-programme academic or social self-efficacy scores. The effects at one year post-programme are sometimes distinguished as “the sustained effect” versus the “the additional follow up effect” to reflect the above difference in the adjustment for just the Time 1 or both the Time 1 and Time 2 outcome measure covariates.

Seeing as data pertaining to career decision self-efficacy and NCEA Level 1 credits and achievement status were only collected at one time point, the lowest possible level of analysis was also across individuals; therefore, these Time 3 outcomes were also assessed using two-level models. Project K and Control groups were compared on each outcome (separately).

The main effects for gender, ethnicity, and school decile rating were also adjusted for in each of the two-level models. Additionally, because a secondary study objective was to determine if the programme was differentially effective for males vs. females, or across different ethnicities (i.e. Māori, Pacific and Asian) and school decile contexts, interactions between each of the aforementioned variables with Condition (i.e. Project K or Control) were assessed for each outcome in the two-level models. Details of the model specifications for these interaction models are also provided below.

**Unconditional Models**

A series of increasingly complex models were estimated to assess programme effects on each outcome beginning with an unconditional intercept-only model (i.e. one with no predictors). For the two-level models, an individual’s outcome score (e.g. academic self-efficacy) is predicted from an intercept (representing the average outcome score for individuals in a particular programme
site) and an error term (individual deviations from the programme site mean of the outcome). The Level-1 intercept (mean outcome for the programme site) is predicted from the grand mean outcome score and error/variance at the programme-level (See Eq. 1.1 & 1.2). A significance test is produced for the Level-2 variance estimate (i.e. \( u_{0j} \) in Eq. 1.2) and this confirms whether or not the mean outcome scores differ significantly across the Level-2 units (e.g. programme sites). In this way, one gains insight as to the likelihood of existing moderating influences at Level-2.

\[
1.1 \quad Y_{ij} = b_{0j} + e_{ij} \\
1.2 \quad b_{0j} = \gamma_{00} + u_{0j}
\]

Calculating the proportion of Level-2 variance to the total variance in the intercept-only model produces an intra-class correlation coefficient (ICC; Hox, 2002) which, in this case, represents the expected correlation between two individuals randomly selected from the same programme site relative to two others randomly selected from different sites. Consequently, this quantifies the degree of dependence in the data and signifies the extent to which contextual factors influence the individual outcome. The ICC was calculated in the above manner for each of the outcomes under investigation in the two-level models.

For the three-level models, the variance explained by programme site-level factors (Level-3) and the variance explained by participant characteristics (Level-2) was calculated by dividing the Level-3 variance by the total model variance (i.e. the sum of Level-1, 2, and 3 variance components for ICC_{Level-3}) and in a separate equation dividing the Level-2 variance by the total model variance (for ICC_{Level-2}).

**Model-Building Process and Specifications**

**Three-Level Models.** The unconditional model for the growth models (M0 in Table 14 for academic self-efficacy and Table 15 for social self-efficacy) was then compared to a linear growth model (M1) which specified Time (pre, post, and one

---

25 Note that the level-1 residual variance for multilevel models with a dichotomous outcome variable (e.g. NCEA Level 1 Achievement Status) is not a constant like that for continuous outcomes but rather an average residual variance across level-2 units (see Ch. 14 of Snijders & Bosker, 1999).
year post-programme) as the only predictor. This provided an overview of the change in academic and social self-efficacy over the three time points across all participants and all programme sites. A similar model (M1a) was produced with both Time and Time\(^2\) to assess the average curvilinear trajectory (as a second polynomial function). Then a model (M2) with the main effects for programme condition, gender, the three ethnicity variables, decile rating, and interactions between each of these variables with Time was produced to assess whether the linear trajectories of academic and social self-efficacy from baseline to one year post-programme was different for Project K vs. Control students, adjusting for the demographic variables. The same model (M2a) was produced for the curvilinear trajectory but an additional interaction between programme condition and the quadratic effect was produced to determine if the quadratic effect was different across programme conditions.

The equation specifications for the final M2 and M2a models are provided below (using academic self-efficacy as the example outcome):

At Level-1, the outcome (ASE or SSE) for time \(t\) for individual \(i\) in programme site \(j\) is predicted from an intercept which represents the outcome for that individual at baseline (\(\pi_{0ij}\)); a slope representing the linear relationship between Time (pre = 0, post = 1, and one year post-programme = 2, entered uncentred) and the outcome; and an error term representing the level-1 residual variance (Eq. 2.1). The curvilinear model specification was similar but included an additional quadratic term for Time (Eq. 2.1a).

\[
\text{2.1} \quad \text{ASE}_{tij} = \pi_{0ij} + \pi_{1ij}(\text{Time}) + e
\]

\[
\text{2.1a} \quad \text{ASE}_{tij} = \pi_{0ij} + \pi_{1ij}(\text{Time}) + \pi_{2ij}(\text{Time}^2) + e
\]

At Level-2, the Level-1 intercept or baseline measure of the outcome for individual \(i\) in programme site \(j\) is predicted from the Level-2 intercept which reflects the average baseline score for the outcome of interest (ASE or SSE) for programme site \(j\); a slope representing the effect of Condition (Control = -0.5, Project K = 0.5, entered uncentered) on the average baseline outcome score for programme site \(j\); slopes representing the effect of Gender (-0.5 = Female, 0.5 =
Male, entered uncentred) and Māori (-0.50 = No, 0.5 = Yes, entered uncentred); Pacific (-0.5 = No, 0.5 = Yes, entered uncentred); and Asian (-0.5 = No, 0.5 = Yes) ethnicity on the baseline score; and an error term representing the residual variance in the predicted intercept at the individual level (Level-2) (Eq. 2.2).

The parameter estimates of primary interest are represented by Eq. 2.3. The first Level-1 slope which represents the growth trajectory of the outcome for individual \( i \) in programme site \( j \) is predicted from the average growth trajectory of the outcome for programme site \( j \); a cross-level interaction representing the effect of Condition on the growth trajectory; a series of cross-level interactions representing the effects of Gender and each of the three ethnicity variables on the growth trajectories; and finally a Level-2 error term representing the residual variance in the predicted slopes (indicating that these slopes were allowed to vary across individual participants).

2.2 \( \pi_{0ij} = b_{00j} + b_{01j}(\text{Condition}_{ij}) + b_{02j}(\text{Gender}_{ij}) + b_{03j}(\text{Māori}_{ij}) + b_{04j}(\text{Pacific}_{ij}) + b_{05j}(\text{Asian}_{ij}) + r_0 \)

2.3 \( \pi_{1ij} = b_{10j} + b_{11j}(\text{Condition}_{ij}) + b_{12j}(\text{Gender}_{ij}) + b_{13j}(\text{Māori}_{ij}) + b_{14j}(\text{Pacific}_{ij}) + b_{15j}(\text{Asian}_{ij}) + r_1 \)

For the M2a models only, a second individual-level slope representing the curvilinear effect for individual \( i \) in programme site \( j \) was also predicted from the average curvilinear effect across all individuals and a slope representing a cross-level interaction between the curvilinear effect and Condition. As stated, this was specified to determine if the curvilinear effect differed between the control and Project K participants (see Eq. 2.4a). The predicted slope was fixed because the model would not converge when the slope was allowed to vary across individual participants.

2.4a \( \pi_{2ij} = b_{20j} + b_{21j}(\text{Condition}_{ij}) \)

School decile rating (Decile) was the only Level-3 predictor included in the model. The main effect of decile on the average baseline outcome score for an individual in programme site \( j \) is predicted from the overall baseline outcome score across all programme sites (the grand mean), a slope representing a cross-level interaction between Decile (entered grand mean centred) and the mean baseline score across programme sites and an error term representing the residual variance.
in the intercept at Level-3 (see Eq. 2.5). Eq. 3.1 specifies the effect of Decile (grand mean centred) on the individual-level growth trajectories. Eqs. 2.6-3.0 represent the grand mean effects of Condition, gender, and the three ethnicity variables, respectively, on the individual-level intercepts (the baseline outcome scores). Eqs. 3.2 – 3.6 represent the grand mean effects of Condition, gender, and the three ethnicity variables on the individual-level slopes for these variables. The $u$’s in equations 2.5 – 3.6 indicate that the intercept and slopes were allowed to vary randomly across programmes. The final two equations (Eqs. 3.7 and 3.8) are only relevant for the M2a models because they specify the prediction of a curvilinear effect ($\text{Time}^2$) for an individual $i$ in programme site $j$ and a cross-level interaction between Condition and the curvilinear effect ($\text{Condition.}\text{Time}^2$). The slopes for the latter two equations were also fixed due to problems of non-convergence when they were allowed to vary.

2.5 $b_{0j} = \gamma_{000} + \gamma_{001}(\text{DECILE}) + u_{00}$
2.6 $b_{01j} = \gamma_{010} + u_{01}$
2.7 $b_{02j} = \gamma_{020} + u_{02}$
2.8 $b_{03j} = \gamma_{030} + u_{03}$
2.9 $b_{04j} = \gamma_{040} + u_{04}$
3.0 $b_{05j} = \gamma_{050} + u_{05}$
3.1 $b_{10j} = \gamma_{100} + \gamma_{101}(\text{DECILE}) + u_{10}$
3.2 $b_{11j} = \gamma_{110} + u_{11}$
3.3 $b_{12j} = \gamma_{120} + u_{12}$
3.4 $b_{13j} = \gamma_{130} + u_{13}$
3.5 $b_{14j} = \gamma_{140} + u_{14}$
3.6 $b_{15j} = \gamma_{150} + u_{15}$
3.7 $b_{20j} = \gamma_{200}$
3.8 $b_{21j} = \gamma_{210}$

**Two Level Models.** For the two-level models, the unconditional model (M0) was then compared to one with the effect of programme condition and the main effects for the demographic predictors (M1); finally a full model (M3) which
includes the previously specified variables and each of the interaction terms (i.e. the product of each demographic variable with Condition) was investigated. Note that when describing the main effect of the programme on each outcome variable I report parameter estimates for the M1 models. In a separate section I discuss the differential effectiveness of the programme for different demographic groups using the coefficients from the M3 models. In this section I comment on any differences obtained in the main programme effect as a result of including interaction terms.

*Model Specifications for Time 2 Self-Efficacy Outcomes*

Immediate post-programme effects for academic and social self-efficacy were evaluated using equations which specified Condition at Level-1 predicting academic or social self-efficacy scores (ASE1/SSE1) at post-programme (Time 2) while statistically adjusting for the following Level-1 covariates: baseline academic or social self-efficacy (ASE0/SSE0, grand mean centred); Gender; Māori, Pacific; and Asian ethnicities (these variables and the Condition variable were contrast-coded and entered uncentred, as described above for the three-level models). As I just explained, Level-1 interaction terms between Condition and Gender, as well as between Condition and each of the three ethnicity variables, were created by multiplying the uncentred Condition and Gender (Cond.Gender), Condition and Māori (Cond.Māori), Condition and Pacific (Cond.Pacific), and Condition and Asian (Cond.Asian) variables together. These were then entered (uncentred) into the Level-1 regression equation (See Eq. 4.0 with Time 2 academic self-efficacy as the example outcome).

\[
4.0 \quad \text{ASE1}_{ij} = b_{0j} + b_{1j}(\text{Condition}_{ij}) + b_{2j}(\text{ASE0}_{ij}) + b_{3j}(\text{Gender}_{ij}) + b_{4j}(\text{Māori}_{ij}) + b_{5j}(\text{Pacific}_{ij}) + b_{6j}(\text{Asian}_{ij}) + b_{7j}(\text{Condition.Gender}_{ij}) + b_{8j}(\text{Condition.Māori}_{ij}) + b_{9j}(\text{Condition.Pacific}_{ij}) + b_{10j}(\text{Condition.Asian}_{ij}) + e_{ij}
\]

At Level-2, the programme site intercept and the slope(s) (representing the relationship between each predictor/covariate and the outcome entered in the Level-1 equation) become predicted outcomes. The grand mean intercept and grand mean slopes of the sample plus programme site-level error (i.e. deviations of programme site mean scores from the grand mean score) are used to calculate
these parameters (Hox, 2002; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). School decile rating (Decile), a Level-2 predictor, was then grand mean centred and added to the Level 2 equations as was a cross-level interaction between Decile (grand mean centred) and Condition (uncentred). Eqs. 4.1 – 5.0 represent the Level-2 equations for the M3 model

\[ b_0 = \gamma_{00} + \gamma_{01}(\text{DECILE}) + u_{0j} \]
\[ b_1 = \gamma_{10} + \gamma_{11}(\text{DECILE}) + u_{1j} \]
\[ b_2 = \gamma_{20} + u_{2j} \]
\[ b_3 = \gamma_{30} + u_{3j} \]
\[ b_4 = \gamma_{40} + u_{4j} \]
\[ b_5 = \gamma_{50} + u_{5j} \]
\[ b_6 = \gamma_{60} + u_{6j} \]
\[ b_7 = \gamma_{70} + u_{7j} \]
\[ b_8 = \gamma_{80} + u_{8j} \]
\[ b_9 = \gamma_{90} + u_{9j} \]

Although this example outlines the model specifications for Time 2 academic self-efficacy as the outcome, the specifications for the Time 2 social self-efficacy outcome were identical with the exception of an exchange of the ASE1 outcome variable with SSE1 and the ASE0 covariate with SSE0.

**Model Specifications for Time 3 Cross-Sectional Outcomes**

Programme effects on career decision self-efficacy (CDSE), NCEA Level 1 credits (NCEACreds), and NCEA Level 1 achievement status (NCEAStat) at the one year post-programme follow up were also investigated with two-level hierarchical linear models. The models for career decision self-efficacy and NCEA Level 1 credit outcomes were specified much like those described above for academic self-efficacy, the exception being there was no need to adjust for a pre-programme measurement of the outcomes of interest at Level-1 because these were not available.
Because the NCEA Level 1 achievement status outcome (NCEAStat) had a dichotomous response format, a Bernoulli hierarchical generalized linear model (HGLM) was specified to evaluate programme effects on the probability of obtaining an NCEA Level 1 certificate (Eq. 6.0). To be specific, Snijders and Bosker (1999) explain that because the parameter estimate for a probability value is only interpretable if it exists within the interval between 0 and 1, the predicted value must first be transformed (Eq. 6.1). In this case, the logit or log of the odds of obtaining an NCEA Level 1 certificate becomes the transformed predicted value ($\eta_{ij}$) which links the structural model including Level-1 predictors to predicted outcome (Eq. 6.2). The Level-1 residual term is not included in the structural model because it is inherently included in the probability equation for the outcome (and the variance for the Level-1 residual term is given by Eq. 6.3). The log-odds can finally be converted to a probability using Eq. 6.4. The Level-2 models followed the same specifications as those for career decision self-efficacy and NCEA Level 1 Credits.

6.0  \[ \text{Prob (NCEAStat} = 1/ b_j) = \varphi_{ij} \]

6.1  \[ \log [\varphi_{ij} / (1 - \varphi_{ij})] = \eta_{ij} \]

6.2  \[ \eta_{ij} = b_0 + b_1(\text{Condition}_{ij}) + b_2(\text{ASE0}_{ij}) + b_3(\text{Gender}_{ij}) + b_4(\text{Māori}_{ij}) + b_5(\text{Pacific}_{ij}) + b_6(\text{Asian}_{ij}) + b_7(\text{Condition.Gender}_{ij}) + b_8(\text{Condition.Māori}_{ij}) + b_9(\text{Condition.Pacific}_{ij}) + b_{10}(\text{Condition.Asian}_{ij}) \]

6.3  \[ \text{var (R}_{ij} = \varphi_{ij}(1- \varphi_{ij}) \]

6.4  \[ \varphi_{ij} = e^{\eta_{ij}} / (1 + e^{\eta_{ij}}) \]

The HLM output produces both population-average and unit-specific parameter estimates for HGLM model. The population-average estimates correct for deviations in the normal distribution of the random effects in the model. Depending on the amount of variance and degree of skew, the estimates will be pulled towards a probability of 0.50. The unit-specific estimates incorporate more information about the distribution of the level-2 units but are more sensitive to model assumptions. Additionally, the latter estimates reflect effects occurring within each Level-2 unit (Raudenbush & Bryk, 2002). For the unit-specific estimates in the current study, the main effect of Condition (for example) on the probability
of obtaining and NCEA Level 1 certificate would be estimated adjusting for programme site (Level-2 units) whereas the population-average estimates would produce an effect of Condition on NCEA Level Status without accounting for differences across sites. As I expect there to be some variation in school achievement outcomes depending on the location of the school/programme site, I report the unit-specific estimates.

Estimates for the Variance Components

The model specifications above indicate that all slopes were allowed to vary, other than the slopes including the curvilinear effect at Level 2 and 3. There was no theoretical justification to presume that the effects would be similar across different programmes or different participants. Indeed the Project K theory of change results (see Chapter Four) suggested that participants respond differently to the programme and different facilitators play a large role in how effective the programmes are. However, the estimates for the random effects tend to be biased when the sample size at Level 2 (or 3) is less than 50 (Maas & Hox, 2004; McCoach, 2010). Of the 50 programme sites in the current set of analyses, few had sufficient data to generate estimates for the variance components for the models of primary interest (M1 and M2). Consequently, other than for the unconditional models (see Table 10), the random effects are not reported for any outcome.

Estimation Procedures

All multilevel analyses were conducted using HLM 6.03 software (Raudenbush, Bryk, & Congdon, 2004). For all of the continuous outcomes (i.e. academic, social, and career decision self-efficacy, and NCEA Level 1 credits), Restricted maximum likelihood estimation (REML) was used. The fixed effects for the HGLM model were estimated using Penalised Quasi Likelihood (PQL) procedures.

Simple Slope Analyses for Interaction Effects

Interaction effects, when significant, should be further investigated to establish where the differences between the slopes lie. When using a regression approach (whether standard or multilevel), interaction effects can be probed using
simple slope analyses (Aiken & West, 1991; Cohen, Cohen, West & Aiken, 2003; Sibley, 2008). The test of the effect for each simple slope may be thought of as roughly equivalent to an independent groups \textit{t-test}. Interaction effects between a binomial predictor and a binomial moderator (e.g. those between Condition and Gender or Condition and any of the ethnicity variables) can be interpreted much like those found in an Analysis of Variance (ANOVA) and the associated simple slopes for these types of interactions represent differences between group means.

All significant (and marginally significant) interactions from the current analyses were further probed using Sibley’s (2008) \textit{Utilities for examining interactions in multiple regression}. These utilities calculate effects for the simple slopes that produce the interaction effect. Accordingly, the utilities can generate mean values for two (user-specified) levels of the predictor for two different levels of the moderator and estimate the significance of the effect for each level of the moderator across the levels of the predictor. The difference between the two levels of the moderator is also calculated for each level of the predictor.

\textit{Model Assumptions}

Because HLM accounts for the dependency in the observations across Level-1 (or Level-2) units when generating parameter estimates, it inherently addresses violations to the assumption of independence that should be considered in ordinary least squares regression. Heterogeneity in the slopes across individuals or groups can also be modelled by allowing the slopes to vary; thus the assumption of homogeneity of variance can also be easily addressed (Field, 2009). The models do however assume normal distributions in the observations of the fixed, as well as the random effects (Hox, 2002; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999).

In Tables 8-9 of the supplementary analyses that form Appendix E, one can obtain the skewness and kurtosis values and their associated Z scores as well as the results from the Kolmogorov-Smirnov (K-S) test assessing violations to the assumption of normality for all of the Level-1 continuous variables included in the analyses (both over the full sample and separately for the Project K and the Control conditions). These results are also reported for the distribution of school decile.
ratings (the only Level-2/3 variable) across the full sample. The results (and visual inspection of the associated histograms) demonstrate that some variables are slightly non-normal; though not enough to merit transformation of the scores because the sample sizes are large at Level-1 (and Level-2 for the growth models). With large samples even negligible deviations from normality generate a significant K-S value (Field, 2009). However, with only 50 units at Level-2, the significant K-S value for the Decile variable is more concerning.

For the above reasons and because the estimates for the fixed effects are influenced by the distribution of the random effects at each level, results for the multilevel models are reported for estimates calculated with robust standard errors as these are less sensitive to departures from normality (Raudenbush & Bryk, 2002). Raudenbush and Bryk (2002) explain that large differences in the estimates of the robust standard errors to those from the standard model should serve as a warning that random effects may be misspecified. Similarities across the models suggest the inferences made about the estimates should be invariant to any deviations from distributional assumptions. Accordingly, these were contrasted for the current analyses and I note (below the relevant tables) where any discrepancies in the significance of the effect occurred.

Slopes across programme sites and across individuals were also graphed for each outcome to ascertain if there were any outlying Level-2 units. A major outlying programme site was identified when examining the Level-2 slopes for NCEA Credits. One programme site reported an average of 52 more NCEA Level 1 credits than any other programme. As it was an extreme outlier, I present the results of for NCEA Level 1 credits with this programme site removed.

RESULTS

DESCRIPTIVE STATISTICS AND BIVARIATE CORRELATIONS

Table 11 presents the means and standard deviations for each of the outcome variables included in the above-specified models. The standardised mean-difference effect sizes (i.e. Cohen’s $d$), along with the confidence interval for
Project K’s Effectiveness

With the exception of baseline academic and social self-efficacy scores, Project K participants reported higher mean levels of self-efficacy for each domain under investigation as well as a few more NCEA Level 1 credits than Control participants. Additionally, a slightly higher proportion of Project K participants obtained an NCEA Level 1 certificate compared to their Control counterparts. For both conditions, the lowest levels of social self-efficacy were obtained at baseline, followed by post-programme, and the highest levels were observed one year after programme completion, which is suggestive of an increasing trend over time. A similar pattern was observed for academic self-efficacy when reported by Control participants; however, even though the Project K participants also reported the lowest levels of academic self-efficacy at baseline, their levels of academic self-efficacy were higher at the end of the programme than one year after programme completion. This is indicative of a curvilinear trend for Project K participants with academic self-efficacy slowly diminishing over time but not declining to its original levels.

According to early guidelines established by Cohen (as cited in DuBois et al., 2011), the effect sizes representing the pre to post-programme effects for academic and social self-efficacy are moderate in size (~ .50), the pre to one-year post-programme effects are small to moderate (i.e. between .20 and .50), and the additional follow up effect of the programme between post and one year post-programme is virtually non-existent for academic self-efficacy and for social self-efficacy. The effect between the two groups on career decision self-efficacy is also
small to moderate but negligible for both NCEA outcomes. The confidence intervals for the academic and social self-efficacy additional follow up effects and the NCEA outcome effects include zero thus this suggests that they are not significant effects.
### Table 11. Descriptive statistics for Level-1 continuous variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Project K</th>
<th>Control</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Means (SD)</td>
<td>n</td>
</tr>
<tr>
<td>Time 1 Academic Self-Efficacy</td>
<td>516</td>
<td>3.26 (.91)</td>
<td>576</td>
</tr>
<tr>
<td>Time 2 Academic Self-Efficacy</td>
<td>471</td>
<td>3.75 (.95)</td>
<td>420</td>
</tr>
<tr>
<td>Time 3 Academic Self-Efficacy</td>
<td>251</td>
<td>3.65 (.92)</td>
<td>236</td>
</tr>
<tr>
<td>Time 1 Social Self-Efficacy</td>
<td>516</td>
<td>3.94 (.88)</td>
<td>576</td>
</tr>
<tr>
<td>Time 2 Social Self-Efficacy</td>
<td>471</td>
<td>4.51 (.78)</td>
<td>420</td>
</tr>
<tr>
<td>Time 3 Social Self-Efficacy</td>
<td>309</td>
<td>4.63 (.76)</td>
<td>285</td>
</tr>
<tr>
<td>Time 3 Career Decision Self-Efficacy</td>
<td>309</td>
<td>3.62 (.72)</td>
<td>284</td>
</tr>
<tr>
<td>Time 3 NCEA Level 1 Credits(^2)</td>
<td>384</td>
<td>89.06 (35.10)</td>
<td>396</td>
</tr>
<tr>
<td>Time 3 NCEA Achievement Status</td>
<td>389</td>
<td>.62 (.49)</td>
<td>401</td>
</tr>
</tbody>
</table>

\(^2\) Descriptives for NCEA Level 1 Credits are reported without the outlying programme site.

### Table 12. Standardised mean difference scores between Project K and Control accounting for baseline scores.

<table>
<thead>
<tr>
<th>Paired Variables</th>
<th>Effect Size</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Self-Efficacy (Time 1 - 2)</td>
<td>.53</td>
<td>[.42 -.64]</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 1 - 3)</td>
<td>.33</td>
<td>[.13 -.53]</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 2 - 3)</td>
<td>-.01</td>
<td>[-.17 -.15]</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 1 - 2)</td>
<td>.49</td>
<td>[-.34 -.63]</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 1 - 3)</td>
<td>.44</td>
<td>[.26-.62]</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 2 - 3)</td>
<td>-.10</td>
<td>[-.26 -.06]</td>
</tr>
</tbody>
</table>
The bivariate correlations (see Table 13) between Condition (Project K or Control) and the outcome variables also provide some support for the effectiveness of the Project K programme as they represent mean differences between the group scores. Being in Project K was associated with having significantly higher mean scores for academic and social self-efficacy upon programme completion as well as one year following programme completion, while there was no significant association between these variables prior to programme start. Mean levels of career-decision self-efficacy were also significantly higher for the Project K group one year following programme completion but no significant associations between Condition and the academic achievement variables (i.e. NCEACreds and NCEAStat) were obtained.

Being female was also associated with greater academic self-efficacy at pre-programme and one year post-programme but no other significant associations with Gender were found. Identifying as Māori was associated with lower levels of academic self-efficacy before the programme and at the end of the programme but this effect was not evident at the one year follow up. Considering the well-supported theoretical association between self-efficacy and performance, it was not surprising that this pattern of associations was mirrored with the academic achievement outcomes. In contrast, Māori ethnicity was associated with significantly higher levels of social self-efficacy across the three time waves. Pacific students had similar results. Significant negative associations between identifying with a Pacific ethnicity and academic self-efficacy were obtained for the baseline measure but not for the post-programme or one year post-programme measures. The number of NCEA Level 1 credits were significantly less for Pacific students and these students were also less likely to have achieved the requirements for obtaining the Level 1 certificate but were more likely to have higher levels of social self-efficacy prior to the programme and at post-programme. Asian participants, conversely, had nearly the reverse trend. Those identifying as Asian were more likely to have higher academic self-efficacy across the three time points, but lower pre and post-programme social self-efficacy in addition to lower career decision self-efficacy one year following the programme. As one would expect, all three
measures of academic self-efficacy were significantly correlated with each other, as were all measures of social self-efficacy.
## Table 13. Bivariate Correlations between Level-1 Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
<th>13.</th>
<th>14.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Time 1 Academic Self-Efficacy</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Time 2 Academic Self-Efficacy</td>
<td>.34**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Time 3 Academic Self-Efficacy</td>
<td>.35**</td>
<td>.63**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Time 1 Social Self-Efficacy</td>
<td>.45**</td>
<td>.03</td>
<td>.09</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Time 2 Social Self-Efficacy</td>
<td>.07*</td>
<td>.43**</td>
<td>.25**</td>
<td>.37**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Time 3 Social Self-Efficacy</td>
<td>.05</td>
<td>.25**</td>
<td>.45**</td>
<td>.35**</td>
<td>.55**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Time 3 Career Decision Self-Efficacy</td>
<td>.10*</td>
<td>.35**</td>
<td>.50**</td>
<td>.16**</td>
<td>.36**</td>
<td>.51**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Time 3 NCEA Level 1 Credits&lt;sup&gt;27&lt;/sup&gt;</td>
<td>.09*</td>
<td>.33**</td>
<td>.29**</td>
<td>-.12**</td>
<td>.08*</td>
<td>.09a</td>
<td>.15**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Time 3 NCEA Achievement Status</td>
<td>.02</td>
<td>.25**</td>
<td>.22**</td>
<td>-.10**</td>
<td>.05</td>
<td>.08</td>
<td>.15**</td>
<td>.73**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Condition</td>
<td>-.01</td>
<td>.22**</td>
<td>.14**</td>
<td>-.04</td>
<td>.22**</td>
<td>.19**</td>
<td>.17**</td>
<td>.04</td>
<td>.05</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Gender</td>
<td>-.06*</td>
<td>-.06</td>
<td>-.12*</td>
<td>.02</td>
<td>-.07</td>
<td>-.07</td>
<td>-.01</td>
<td>-.04</td>
<td>-.04</td>
<td>.02</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Māori</td>
<td>-.06*</td>
<td>-.13**</td>
<td>-.04</td>
<td>.09**</td>
<td>.07*</td>
<td>.10*</td>
<td>-.03</td>
<td>-.21**</td>
<td>-.15**</td>
<td>-.03</td>
<td>-.03</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Pasifika</td>
<td>.20**</td>
<td>.06</td>
<td>-.02</td>
<td>.21**</td>
<td>.08*</td>
<td>.03</td>
<td>-.07</td>
<td>-.27**</td>
<td>-.26**</td>
<td>-.05</td>
<td>-.05</td>
<td>-.03</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>14. Asian</td>
<td>.10**</td>
<td>.10**</td>
<td>.12**</td>
<td>-.08**</td>
<td>-.07*</td>
<td>-.01</td>
<td>-.09*</td>
<td>.04</td>
<td>.06</td>
<td>-.05</td>
<td>-.03</td>
<td>-.06*</td>
<td>.007</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: ** p < .01  
* p < .05  
a p < .10  

<sup>27</sup> The programme site with outlying mean scores on NCEA Level 1 credits is not included in the calculations of the bivariate correlations for NCEA Level 1 credits.
Though the preliminary results are encouraging, to ascertain more definitive conclusions about the effects of Project K on the self-efficacy and academic achievement outcomes the effects must be assessed while accounting for the potential dependency in the observations and while also statistically adjusting for baseline levels of self-efficacy. In order to better understand our secondary research questions of how programme results may differ across participant demographic variables and social contexts (i.e. school decile ratings) it was also necessary to model the interactions of these factors with the programme effects, while adjusting for the main effects of programme condition, gender, ethnicity, and school decile rating. Hence, the multilevel models derived from these analyses are presented next.

**UNCONDITIONAL MODELS AND INTRA-CLASS CORRELATIONS**

Three-Level Models

The unconditional models for the three-level growth trajectory (M0 in Tables 14 and 15) provide little information of interest because the intercept reflects the average academic (Table 14) and social self-efficacy (Table 15) scores across all the three time points; however, the ICCs calculated from the Level 2 and Level 3 variance components (Table 16) represent the expected correlation on each of these outcomes between two of the three time point self-efficacy scores chosen randomly from the same individual relative to two randomly selected scores from a different individual (ICC$_{Level 2}$), and the expected association between two individuals randomly chosen from the same programme site relative to two individuals selected randomly from a different site (ICC$_{Level 3}$). Thus it signifies the degree of dependency across the observations at these levels. The ICC$_{Level 2}$ is 0.35 and 0.31 for academic and social self-efficacy, respectively and the ICC$_{Level 3}$ is 0.04 for both outcomes. All values are based on intercepts that vary significantly across individuals ($r_0$) and across programme sites ($u_0$).

In terms of the fixed coefficients, the unconditional linear and curvilinear trajectory models (M1 and M1a) are much more informative as the intercept represents the baseline scores for academic and social self-efficacy and the
regression coefficient for Time reflects the average rate of change over time. The coefficient for Time$^2$ indicates whether or not the trajectory accelerates or decelerates at any point and to what degree (Singer & Willet, 2003). The results for the M1 models (Tables 14 and 15) demonstrate that, in general, participants began the programme with an academic self-efficacy score of 3.31 ($SE = 0.05$) on a scale from 1 (not well at all) to 6 (quite well) and the linear trajectory increased significantly over time ($b = 0.14, p < 0.001$). The average baseline score for social self-efficacy was a slightly higher 4.01 with a standard error of 0.04 and also increased significantly from pre to one year post-programme ($b = 0.26, p < 0.001$). However, when accounting for the quadratic term, the mean baseline scores were slightly lower ($M_{\text{academic}} = 3.27, SE = 0.05$ and $M_{\text{social}} = 3.98, SE = 0.04$); the linear component of the slope for both outcomes remained significant and increased at a higher rate ($b = 0.46, p < 0.001$ for academic and $b = 0.47, p < 0.001$ for social); and the quadratic effect was significant and negative for each outcome ($b = -0.18, p < 0.001$ for academic and $b = -0.12, p < 0.001$ for social). This indicates that on average the participants’ (across both programme conditions) academic and social self-efficacy scores increased approximately half a response category on the 6-point Likert scale. This growth eventually decelerated but at less than half the rate of the initial growth.

**Two-level Models**

The intercept in the two-level unconditional models (M0 in Tables 18-22) represent the mean for the outcome variable of interest across all participants in all programmes without including any predictors/covariates. For the academic and social self-efficacy outcomes, these values ranged from 3.53 with a standard error of 0.05 (one year post-programme academic) to 4.49 with a standard error of 0.04 (one year post-programme social) on a scale from 1 (not well at all) to 6 (quite well). The mean for career decision self-efficacy was 3.50 ($SE = 0.03$) on scale of 1 (No confidence at all) to 5 (Complete confidence). Finally, the grand mean for the number of NCEA Level 1 credits was 87.56 with a standard error of 2.26 and the probability of obtaining an NCEA Level 1 certificate without considering any other predictors was 0.61.
Table 17 reports the variance components for the intercepts ($u_0$) and the ICCs for the unconditional models of all two-level model outcomes. This demonstrates that there was marginally significant variance in the intercepts for Time 2 and Time 3 academic self-efficacy and Time 3 social self-efficacy. The respective ICCs for these variables were 0.02, 0.03, and 0.03. The intercept for Time 2 social self-efficacy did vary significantly across Level-2 units as did those for NCEA Level 1 credits and Status. The ICC for Time 2 social self-efficacy was 0.04 while those for Time 3 NCEA Level 1 credits and NCEA Level 1 Status were 0.13 and 0.15, respectively. In contrast, the data for career-decision self-efficacy indicated complete independence in the observations as the ICC was 0 and the intercept did not vary significantly across programme sites.

Overall, the results from the unconditional models show that when one does not specify any predictors, 35% and 31% of the unexplained variance in the average academic and social self-efficacy scores across all time points can be attributed to true individual differences plus error at this level, whereas 4% of the remaining variance in both outcomes occurs at the programme site level. When considering academic and social self-efficacy outcomes at post and one year post-programme, a small portion (2-4%) can be attributed to factors within the participants’ shared programme site context (or error at this level) and a sizable portion of the unexplained variance in NCEA Level 1 credits (13%) and NCEA Level 1 Status (15%) could potentially be explained by such factors. However, none of the unexplained variance in the participants’ career decision self-efficacy scores should be attributed to factors at the programme site level. Furthermore, where significant, the variance across intercepts (at either the individual or the programme site level) indicates that the associated ICCs reflect dependency across the observations that is not negligible and thus should be accounted for when estimating other effects on these outcomes.
### Table 14. Multilevel modelling results for academic self-efficacy linear and curvilinear growth trajectories

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Academic Self-Efficacy Linear Trajectory</th>
<th>Academic Self-Efficacy Curvilinear Trajectory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M0</td>
<td>M1</td>
</tr>
<tr>
<td>Intercept</td>
<td>3.41 (0.04)***</td>
<td>3.31(0.05)***</td>
</tr>
<tr>
<td>Time(γ100)</td>
<td>-</td>
<td>0.14 (0.03)***</td>
</tr>
<tr>
<td>Time²(γ200)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Condition(γ010)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender(γ020)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Māori(γ030)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pacific Peoples(γ040)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian(γ050)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Condition.Time (γ110)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Condition.Time² (γ120)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender.Time (γ120)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Māori.Time (γ130)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pacific.Time (γ140)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian.Time (γ150)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile (γ160)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile.Time (γ170)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes. Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses. *** p < .001, ** p < .01, * p < .05, p < .10. Note. The marginal main effect of Gender in M3 was non-significant and the Pacific.Time interaction effect in M3a was marginal when robust standard errors were not used.
### Table 15. Multilevel modelling results for social self-efficacy linear and curvilinear growth trajectories

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Social Self-Efficacy Linear Trajectory</th>
<th>Social Self-Efficacy Curvilinear Trajectory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M0</td>
<td>M1</td>
</tr>
<tr>
<td>Intercept ($\gamma_{000}$)</td>
<td>4.20 (0.03)***</td>
<td>4.01 (0.04)***</td>
</tr>
<tr>
<td>Time($\gamma_{100}$)</td>
<td>-</td>
<td>0.26 (0.02)***</td>
</tr>
<tr>
<td>Time$^2$($\gamma_{200}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Condition($\gamma_{010}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender($\gamma_{020}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Māori($\gamma_{030}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pacific Peoples($\gamma_{040}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian($\gamma_{050}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Condition.Time ($\gamma_{110}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Condition.Time$^2$ ($\gamma_{210}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender.Time ($\gamma_{120}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Māori.Time ($\gamma_{130}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pacific.Time ($\gamma_{140}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian.Time ($\gamma_{150}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile ($\gamma_{160}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile.Time ($\gamma_{170}$)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes. Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses. *** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .10$. Note. The Asian.Time and Decile.Time interaction effects were not significant when robust standard errors were not used.
Table 16. Variance components and intra-class correlation coefficients for three-level unconditional models

<table>
<thead>
<tr>
<th>Variance Components</th>
<th>Academic Self-Efficacy</th>
<th>Social Self-Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_{00}$</td>
<td>0.04***</td>
<td>0.03**</td>
</tr>
<tr>
<td>$R_0$</td>
<td>0.31***</td>
<td>0.24**</td>
</tr>
<tr>
<td>$E$</td>
<td>0.54</td>
<td>0.51</td>
</tr>
<tr>
<td>ICC_{Level 2}</td>
<td>0.35</td>
<td>0.31</td>
</tr>
<tr>
<td>ICC_{Level 3}</td>
<td>0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Notes. ** p. < .01
*p. < .05

Table 17. Variance components and intra-class correlation coefficients for two-level models

<table>
<thead>
<tr>
<th>Outcomes for Two-Level Models</th>
<th>$u_0$</th>
<th>$R$</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time 2 Academic Self-Efficacy</td>
<td>0.02$^a$</td>
<td>0.91</td>
<td>0.02</td>
</tr>
<tr>
<td>Time 3 Academic Self-Efficacy</td>
<td>0.03$^a$</td>
<td>0.83</td>
<td>0.03</td>
</tr>
<tr>
<td>Time 1 Social Self-Efficacy</td>
<td>0.03**</td>
<td>0.70</td>
<td>0.04</td>
</tr>
<tr>
<td>Time 3 Social Self-Efficacy</td>
<td>0.02$^a$</td>
<td>0.66</td>
<td>0.03</td>
</tr>
<tr>
<td>Time 3 Career Decision Self-Efficacy</td>
<td>0.00</td>
<td>0.51</td>
<td>0.00</td>
</tr>
<tr>
<td>Time 3 NCEA Level 1 Credits</td>
<td>161.84**</td>
<td>1116.42</td>
<td>0.13</td>
</tr>
<tr>
<td>Time 3 NCEA Achievement Status</td>
<td>0.57**</td>
<td>-</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Notes. ** p. < .01
*p. < .01
*p < .10.
Table 18. Multilevel model results for post-programme academic and social self-efficacy

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Academic Self-Efficacy</th>
<th>Social Self-Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Time 2)</td>
<td>(Time 2)</td>
</tr>
<tr>
<td>Regression coefficients (fixed effects)</td>
<td>M0</td>
<td>M1</td>
</tr>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>3.55 (0.04)***</td>
<td>3.55 (0.08)***</td>
</tr>
<tr>
<td>Condition ($\gamma_{10}$)</td>
<td>-</td>
<td>0.45 (0.01)***</td>
</tr>
<tr>
<td>Time 1 ASE0/SSE0 ($\gamma_{100}$)</td>
<td>-</td>
<td>0.35 (0.04)***</td>
</tr>
<tr>
<td>Gender ($\gamma_{20}$)</td>
<td>-</td>
<td>-0.06 (0.05)</td>
</tr>
<tr>
<td>Māori ($\gamma_{30}$)</td>
<td>-</td>
<td>-0.24 (0.07)**</td>
</tr>
<tr>
<td>Pacific Peoples ($\gamma_{40}$)</td>
<td>-</td>
<td>0.01 (0.07)</td>
</tr>
<tr>
<td>Asian ($\gamma_{50}$)</td>
<td>-</td>
<td>0.18 (0.15)</td>
</tr>
<tr>
<td>Cond.Gender ($\gamma_{60}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Māori ($\gamma_{70}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Pacific ($\gamma_{80}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Asian ($\gamma_{90}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile ($\gamma_{101}$)</td>
<td>-</td>
<td>0.00 (0.01)</td>
</tr>
<tr>
<td>Decile.Cond ($\gamma_{111}$)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes. Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses. *** p < .001, ** p < .01, *p<0.05, *p < .10
Table 19. Multilevel model results for one year post-programme academic self-efficacy

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Academic Self-Efficacy (Time 3 controlling for ASE0)</th>
<th>Academic Self-Efficacy (Additional Follow Up) (Time 3 controlling for ASE0 and ASE1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression coefficients</td>
<td>M0</td>
<td>M1</td>
</tr>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>3.53 (0.05)**</td>
<td>3.55 (0.10)**</td>
</tr>
<tr>
<td>Condition ($\gamma_{10}$)</td>
<td>-</td>
<td>0.30 (0.08)**</td>
</tr>
<tr>
<td>Time 1 ASE0 ($\gamma_{200}$)</td>
<td>-</td>
<td>0.39 (0.05)**</td>
</tr>
<tr>
<td>Time 2 ASE1 ($\gamma_{300}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender ($\gamma_{20}$)</td>
<td>-</td>
<td>-0.18 (0.09)*</td>
</tr>
<tr>
<td>Māori ($\gamma_{30}$)</td>
<td>-</td>
<td>0.05 (0.11)</td>
</tr>
<tr>
<td>Pacific Peoples ($\gamma_{40}$)</td>
<td>-</td>
<td>-1.41 (0.10)</td>
</tr>
<tr>
<td>Asian ($\gamma_{50}$)</td>
<td>-</td>
<td>0.18 (0.17)</td>
</tr>
<tr>
<td>Cond.Gender ($\gamma_{60}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Māori ($\gamma_{70}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Pacific ($\gamma_{80}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Asian ($\gamma_{90}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile ($\gamma_{10}$)</td>
<td>-</td>
<td>0.00 (0.01)</td>
</tr>
<tr>
<td>Decile.Cond ($\gamma_{11}$)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes. Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses. *** $p < .001$, ** $p < .01$, *$p<0.05$, p < .10. Note. The main effect of Condition in M1 for the additional follow up was only marginal when robust standard errors were not used.
### Table 20. Multilevel model results for one year post-programme social self-efficacy

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Social Self-Efficacy (Time 3 controlling for SSE0)</th>
<th>Social Self-Efficacy (Follow Up) (Time 3 controlling for SSE0 and SSE1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M0</td>
<td>M1</td>
</tr>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>4.49 (0.04)**</td>
<td>4.59 (0.08)**</td>
</tr>
<tr>
<td>Condition ($\gamma_{10}$)</td>
<td>-</td>
<td>0.34 (0.06)**</td>
</tr>
<tr>
<td>Time 1 ASE0 ($\gamma_{100}$)</td>
<td>-</td>
<td>0.33 (0.03)**</td>
</tr>
<tr>
<td>Time 2 ASE1 ($\gamma_{110}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender ($\gamma_{20}$)</td>
<td>-</td>
<td>-0.12 (0.05)*</td>
</tr>
<tr>
<td>Māori ($\gamma_{20}$)</td>
<td>-</td>
<td>0.17 (0.08)*</td>
</tr>
<tr>
<td>Asian ($\gamma_{30}$)</td>
<td>-</td>
<td>0.03 (0.10)</td>
</tr>
<tr>
<td>Cond.Gender ($\gamma_{60}$)</td>
<td>-</td>
<td>-0.07 (0.11)</td>
</tr>
<tr>
<td>Cond.Māori ($\gamma_{70}$)</td>
<td>-</td>
<td>-0.36 (0.13)**</td>
</tr>
<tr>
<td>Cond.Pacific ($\gamma_{80}$)</td>
<td>-</td>
<td>-0.03 (0.16)</td>
</tr>
<tr>
<td>Cond.Asian ($\gamma_{90}$)</td>
<td>-</td>
<td>0.39 (0.19)*</td>
</tr>
<tr>
<td>Decile ($\gamma_{11}$)</td>
<td>-</td>
<td>0.02 (0.01)</td>
</tr>
<tr>
<td>Decile.Cond ($\gamma_{111}$)</td>
<td>-</td>
<td>-0.02 (0.02)</td>
</tr>
</tbody>
</table>

Notes. Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses.*** $p < .001$, ** $p < .01$, * $p < .05$, $p < .10$. Note. For Time 3 social self-efficacy adjusting for Time 1, the main effect of Gender in M1 and M2, and the main effect for Māori in M1 were only marginal and the Cond.Asian interaction in M2 was not significant when robust standard errors were not used. For Time 3 social self-efficacy adjusting for Time 1 and Time 2, the main effect of Gender in M1 and M2, the main effect of Asian in M2, and the Decile.Condition interaction in M2 were not significant and the main effects of Māori in M2 and Decile in M1 were only marginal when robust standard errors were not used.
### Table 21. Multilevel model results for one year post-programme career-decision self-efficacy and NCEA Level 1 Credits

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Career Decision Self-Efficacy (Time 3)</th>
<th>NCEA Level 1 Credits – without Level 2 Outlier (Time 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M0</td>
<td>M1</td>
</tr>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>3.50 (0.03)**</td>
<td>3.37 (0.06)**</td>
</tr>
<tr>
<td>Condition ($\gamma_{10}$)</td>
<td>-</td>
<td>0.23 (0.06)**</td>
</tr>
<tr>
<td>Gender ($\gamma_{20}$)</td>
<td>-</td>
<td>-0.03 (0.06)</td>
</tr>
<tr>
<td>Māori ($\gamma_{30}$)</td>
<td>-</td>
<td>-0.05 (0.07)</td>
</tr>
<tr>
<td>Pacific Peoples ($\gamma_{40}$)</td>
<td>-</td>
<td>-0.06 (0.09)</td>
</tr>
<tr>
<td>Asian ($\gamma_{50}$)</td>
<td>-</td>
<td>-0.22 (0.10)*</td>
</tr>
<tr>
<td>Cond.Gender ($\gamma_{60}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Māori ($\gamma_{70}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Pacific ($\gamma_{80}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Asian ($\gamma_{90}$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Decile ($\gamma_{10}$)</td>
<td>-</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Decile.Cond ($\gamma_{11}$)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Notes.** Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses.*** $p < .001$, ** $p < .01$, * $p < .05$, $\hat{p} < .10$. For career decision self-efficacy, the main effect of Asian in M1 and M2 was only marginal and the Condition.Asian interaction effect was non-significant in M2 when robust standard errors were not used. For NCEA Level 1 credits, the main effect of Condition and of Decile and the Decile.Condition interaction in M2 were all non-significant and the main effect of Gender, Asian, and the Condition.Asian interaction effect in M2 were only marginal when the robust standard errors were not used.
Table 22. Multilevel model results for one year post-programme NCEA Level 1 Credits

<table>
<thead>
<tr>
<th>Parameters</th>
<th>NCEA Level 1 Achievement Status (Time 3) unit-specific models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>0.46 (0.13)**</td>
</tr>
<tr>
<td>Condition ($\gamma_{10}$)</td>
<td>-</td>
</tr>
<tr>
<td>Gender ($\gamma_{20}$)</td>
<td>-</td>
</tr>
<tr>
<td>Māori ($\gamma_{30}$)</td>
<td>-</td>
</tr>
<tr>
<td>Pacific Peoples ($\gamma_{40}$)</td>
<td>-</td>
</tr>
<tr>
<td>Asian ($\gamma_{50}$)</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Gender ($\gamma_{60}$)</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Māori ($\gamma_{70}$)</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Pacific ($\gamma_{80}$)</td>
<td>-</td>
</tr>
<tr>
<td>Cond.Asian ($\gamma_{90}$)</td>
<td>-</td>
</tr>
<tr>
<td>Decile ($\gamma_{101}$)</td>
<td>-</td>
</tr>
<tr>
<td>Decile.Cond ($\gamma_{111}$)</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes. Estimates of fixed effects calculated with robust standard errors. Parameter estimate standard errors listed in parentheses. *** p < .001, ** p < .01, *p<0.05, *p < .10. The marginal main effect of Decile was non-significant when robust standard errors were not used.
MAIN EFFECT OF PROJECT K ON DESIRED OUTCOMES

Three-Level Models

Linear Trajectories

The non-significant main effect of Condition on the intercepts in the academic and social self-efficacy growth models indicates that there was no difference between Project K and Control in the baseline levels of these measures (see Tables 14 and 15). However, significant cross-level interactions between Condition and Time for both academic ($b = 0.25, p < .001$) and social self-efficacy ($b = 0.21, p < .001$) in the M3 models indicate that the linear trajectory for these scores (from baseline to one year post-programme) differ depending on programme condition.

The simple slopes analyses for academic self-efficacy revealed that the linear increase from pre to one year post-programme was significant for the Project K group (simple slope = 0.19, $t = 4.54, p < .001$) but not the Control group (simple slope = -0.06, $t = -1.56, p = .12$). Confirming the above finding for the intercept, there was no significant difference between Project K and Control in their mean baseline scores ($M_{PK} = 3.56$ and $M_{Control} = 3.61$) but there was at one year post-programme. Project K participants reported higher levels than the Control group ($M_{PK} = 3.99; M_{Control} = 3.45$; simple slope = 0.54, $t = 6.22, p < .001$). Figure 8 presents the slopes for each group.

![Figure 8](image-url)

_Figure 8. Linear growth trajectory of academic self-efficacy for Project K and Control from baseline to one year post-programme._
The simple slopes analyses for social self-efficacy demonstrated that the linear trajectory for both Project K and Control was significant over the three time points (simple slope = 0.15, t = 3.80, \( p < .001 \) for Control; simple slope = 0.36, \( t = 8.77, p < .001 \) for Project K). As with academic self-efficacy, there was no significant difference in the means between the groups at baseline (\( M_{PK} = 4.11 \) and \( M_{Control} = 4.10 \)) however, they were significantly different one year after programme completion. Again, Project K participants reported higher scores (\( M_{PK} = 4.82; M_{Control} = 4.40; \) simple slope = 0.42, \( t = 4.94, p < .001 \)). Figure 9 depicts this relationship.

Figure 9. Linear growth trajectory of social self-efficacy for Project K and Control from baseline to one year post-programme.

Curvilinear Effects

The significant curvilinear effect obtained across participants discussed earlier for the M1a models suggested that the linear effect alone does not accurately capture what is happening for participants during this period for academic and social self-efficacy outcomes. Furthermore, when the quadratic term (\( T_{ime}^2 \)) and a cross-level interaction between Condition and \( T_{ime}^2 \) were added to the M3 model (refer to column M3a in Tables 14 and 15), the significant cross-level interaction effect between Condition and \( T_{ime} \) (i.e. the linear trajectory) increased substantially for both social (\( b = 0.68, p < 0.001 \)) and academic self-efficacy (\( b = 0.73, p < .001 \)). Moreover, the cross-level interaction between Condition and \( T_{ime}^2 \) was significant but negative for both outcomes (\( b = -0.27, p < .001 \) for academic and
b = -0.25, p < .001 for social). This indicates that the rate of deceleration we obtained for these outcomes across all individuals in M1a differed depending on programme condition.

Sibley’s (2008) Utilities for Examining Multiple Regression Interactions provides a calculator for curvilinear interactions which generates values for the linear and curvilinear effects across two levels of the moderator (e.g. Project K or Control) and the estimated means for these levels at several user-specified time points. Using this utility the Project K group had an estimated linear effect of 0.75 and a curvilinear effect of -0.32 for academic self-efficacy. Additionally this group had an estimated baseline academic self-efficacy score of 3.50, a post-programme score of 3.93, and a one year post-programme score of 3.73. The Control group had an estimated linear effect of 0.02 and a curvilinear effect of -0.05 for this outcome and associated scores of 3.59, 3.56, and 3.44 at baseline, post-programme, and one year post-programme, respectively (See Figure 10).

![Figure 10. Curvilinear growth trajectory of academic self-efficacy for Project K and Control from baseline to one year post-programme.](image)

For social self-efficacy the estimated means for the Project K group were 4.05, 4.62, and 4.72 over the three time points; the linear effect was 0.80 and the curvilinear effect was -0.24. For Control participants baseline scores were 4.11 on average, followed by 4.25 at post-programme and 4.41 one year after Project K had completed. The linear effect was estimated to be 0.12 and the curvilinear effect was 0.02 (See Figure 11).
Two-Level Main Effect Models (M1)

The following analyses provide estimates for the main effect of Condition on each outcome in the two-level models (all based on the M1 models in Tables 18-22). Because utilities do not exist to test the significance of the differences between groups at different points of a curvilinear trajectory the analyses of the two level models comparing Project K and the Control groups on academic and social self-efficacy at the immediate end of the programme and one year post-programme provide a more detailed picture of group differences at these particular points in time. Beginning with immediate post-programme effects, Project K participants on average had significantly higher academic (b = 0.45, \( p < .001 \)) and social self-efficacy scores (b = 0.38, \( p < .001 \)) than Control participants adjusting for the baseline level of each outcome and the main effects of gender, the ethnic variables, and school decile rating (see Table 18).

Tables 19 and 20 depict what is happening for academic and social self-efficacy outcomes one year after programme completion. With regards to the sustained programme effect (adjusting for baseline levels in the outcome and demographic variables), the results demonstrated that, on average, Project K participants still had higher levels of academic (b = 0.30, \( p < .001 \)) and social (b = 0.34, \( p < .001 \)) self-efficacy than their Control counterparts. Interestingly, Condition also had an additional follow up effect on social self-efficacy (b = 0.17, \( p < .01 \)) and
academic self-efficacy (b = 0.14, p < .05) one year post-programme (i.e. controlling for both baseline and post-programme levels of the relevant outcome). The influence of Project K participation on career decision self-efficacy one year after programme completion was also significant (b = 0.23, p < .001) while adjusting for the demographic covariates (see Table 21).

Project K participants obtained 3.27 more NCEA Level 1 credits than Control participants on average but this difference was not significant. Likewise, the probability of being awarded an NCEA Level 1 Certificate was 53% for the Project K group and 47% for the Control group but this difference was not significant either (see Table 22 for log-odds and odds-ratios).

**MAIN EFFECTS OF DEMOGRAPHIC VARIABLES ON OUTCOMES**

The main effects for gender, the three ethnicity variables and decile rating for the academic and social self-efficacy are reported for all two-level models (column M1 of Tables 18-22). I do not report the effects of the demographic variables on the growth trajectories as these models were produced to provide an overall picture of what is happening between Project K and Control over time (the main effects for the demographic variables and the interactions between these variables with Time are presented in Tables 14 and 15 for those who are interested). The two-level models were generated to further probe the differences between Project K and Control and also to investigate interactions between programme condition and the demographic variables. It is thus helpful to first have a look at how the demographic variables operated in terms of main effects.

**Gender Differences**

A significant gender difference was obtained for post-programme social self-efficacy when scores were adjusted for baseline social self-efficacy levels, programme condition and the other demographic covariates (b = -0.13, p < .01) and the difference was also significant one year post-programme (b = -0.12, p < .05). The difference was only marginally significant at one year post-programme when assessing the additional follow up effect (b = -0.09, p = .07). Similarly, marginal gender differences were obtained for academic self-efficacy at one year post
programme when controlling for baseline scores, programme condition, and the demographic covariates (\(b = -0.18, p = .051\)) and the difference was significant when evaluating the additional follow up effect of gender (accounting for both baseline and post-programme academic self-efficacy; \(b = -0.15, p < .05\)). In all instances females reported the higher scores.

**Ethnic differences**

Because all ethnicities were contrast-coded (-0.5 and 0.5) with the ethnic group of interest taking the upper value and because each of these were entered into the regression equations simultaneously, the effects for any particular group reflect the effect of that particular ethnicity in comparison to a contrast group of all other individuals (including the non-coded reference group of individuals identifying as European and “Other”).

**Māori Ethnicity**

When investigating post-programme differences, the findings revealed that Māori had lower academic self-efficacy (\(b = -0.24, p <.001\)) but there was no difference between Māori and the contrast group in post-programme social self-efficacy. For the sustained one year post-programme effect, Māori youth had higher levels of social self-efficacy (\(b = 0.17, p <.05\) but the effect was only marginal for the additional social self-efficacy follow up effect (\(b = 0.15, p = .08\)). Māori individuals received fewer NCEA Level 1 Credits (\(b = -17.56, p < .001\)). Accordingly, they also had a lower log-odds of being awarded a Level 1 certificate (\(b = -0.75, p < .01\)). Māori had an expected 40% probability of obtaining the certificate compared to 59% for the contrast group.

**Pacific Ethnicity**

People identifying with a Pacific nation ethnicity had a lower mean number of NCEA Level 1 Credits (\(b = -22.74, p < .001\)) and were less likely to obtain an NCEA Level 1 certificate (\(b = -1.28, p <.001\)). Specifically, 35% of Pacific participants were expected to obtain a certificate whereas those in the contrast group had an expected probability of 66%.
Asian Ethnicity

With regards to career decision self-efficacy one year post-programme, Asians had lower levels ($b = -0.22, p < .05$). They did, however, receive a greater number of NCEA Level 1 credits, on average ($b = 8.74, p < .05$).

School Decile Rating

Interestingly, school decile rating only had a significant main effect on social self-efficacy for the additional one year post-programme follow up effect ($b = 0.02, p < .05$). This indicates participants in higher decile schools were more likely to report additional increases in social self-efficacy between the end of the programme and one year later.

MODERATING INFLUENCES OF DEMOGRAPHIC VARIABLES ON PROJECT K EFFECTS

To address the secondary aim of the study, which was to investigate if Project K was differentially effective for different participant subgroups based on gender, ethnicity, and school decile rating, the interactions between programme condition and these variables were explored. The interaction effects between Condition and each of the demographic variables (i.e. Gender; Māori, Pacific, and Asian ethnicity; and Decile) for each of the two-level outcomes are reported in column M2 in Tables 18-22. Where interactions did exist, simple slope analyses were conducted. Again, when the moderators are binomial (which they all are except Decile), the simple slopes represent differences between the groups and the test of the simple slope is in essence a test of the difference in group means. The main effects for Condition and the demographic variables in these models are reported prior to the interaction effects to provide a comparison with the findings obtained in the M1 models.

Academic Self-Efficacy

For post-programme academic self-efficacy, the main effects for Condition ($b = 0.41, p < .01$) and Māori ethnicity ($b = -0.23, p < .01$) remained significant when all interaction effects were included in the model. Furthermore, none of the interactions were significant.
With regards to one year post-programme academic self-efficacy effect, the main effect of Condition was no longer significant for the sustained or the additional follow up effect. This suggests that the main Condition effects reported previously for the M1 models are accounted for by one or more of the interactions. In contrast, the main effects for gender were significant (b = -0.19, \( p < .05 \) for the sustained effect; b = -0.15, \( p < .05 \) for the additional follow up effect). Additionally, a significant interaction between Condition and Gender was obtained for one year post-programme academic self-efficacy for the sustained effect (b = 0.32, \( p < .05 \)) and a marginally significant interaction was obtained for the additional follow up effect (b = 0.23, \( p < .07 \)).

Simple slope analyses probed these effects further. For the former effect, the difference between females in the Control and Project K groups one year post-programme was not significant. Females in the Control group reported a mean score of 3.64 while those in Project K scored 3.69 on average. However, for males the difference between the groups was significant (simple slope = 0.38, \( t = 2.14, p < .05 \)). Males in the Control group scored 3.28 on average while those in Project K reported a higher mean level of 3.66. Notably, there was no significant gender difference in the mean scores for the Project K group, while there was for Control participants (simple slope = -0.35, \( t = -3.09, p < .001 \)). Figure 12 illustrates these differences.

![Figure 12. Interaction between programme condition and gender for academic self-efficacy one year post-programme (sustained programme effect).](image)
The simple slope analyses for the marginally significant interaction between Condition and Gender for the additional academic self-efficacy follow up effect revealed no differences between Project K and Control group scores for females (\(M_{PK} = 3.64; M_{Control} = 3.67\)) nor for males (\(M_{PK} = 3.60; M_{Control} = 3.40\)). That said, females had higher academic self-efficacy than males in the Control group (simple slope = -0.27, \(t = -2.73, p < .001\)) but no gender difference was obtained in the Project K group (see Figure 13).

![Figure 13](image-url)

**Figure 13.** Interaction between programme condition and gender for the additional academic self-efficacy follow up effect.

**Social Self-Efficacy**

The main effects for Condition (\(b = 0.38, p < .01\)) and Gender (\(b = -0.13, p < .001\)) remained significant when all interaction effects were included in the model for post-programme social self-efficacy. Like post-programme academic self-efficacy, none of the interactions were significant.

When all interactions were included, the main effect of Condition was still significant for one year post-programme social self-efficacy for the sustained effect (\(b = 0.38, p < .01\)) but not for the additional follow up effect. Thus, the additional influence of Condition on this outcome, over and above its influence on post-programme levels, appears to be accounted for by one or more of the interaction
effects. The main effect for Gender was still significant when adjusting for the sustained effect \((b = -0.11, p <.05)\) and it was still marginally significant for the additional follow up effect \((b = -0.10, p = .07)\). Furthermore, the main effect of Māori ethnicity remained significant for the sustained effect \((b = 0.18, p <.05)\) and the previously marginal effect became significant when assessing the additional follow up effect \((b = 0.17, p <.05)\). For the latter model, Asian ethnicity also had a main effect that did not exist in previous models \((b = 0.18, p <.05)\). Decile also had an effect \((b = 0.03, p <.05)\), although this evident in the previous M1 model.

A significant Condition.Māori interaction was obtained for social self-efficacy one year post-programme for the sustained effect \((b = -0.36, p <.01)\). The simple slope analyses demonstrated that there was no difference between Project K and Control for young Māori \((M_{PK} = 4.79, M_{Control} = 4.59)\). There was, however, a significant difference between the groups for the contrast group with Project K participants reporting a higher mean score \((M_{PK} = 4.79, M_{Control} = 4.23; \text{simple slope} = 0.56, t = 4.13, p <.001)\). Further, within the Control group there was a difference between Māori and the contrast group \((\text{simple slope} = 0.36, t = 4.14, p <.001)\) but this difference did not exist within the Project K group (see Figure 14).

![Figure 14](image-url)
Similarly, a significant interaction between programme condition and Māori ethnicity (Cond.Māori) was obtained for the additional social self-efficacy follow up effect \( (b = -0.41, p < .01) \). Figure 15 illustrates the interaction. Simple slope analyses showed a difference between Project K and Control in the contrast group \( (\text{simple slope} = 0.25, t = 2.11, p < .05) \) with the Control group reporting a mean score of 4.43 and Project K a mean of 4.68. Again, there was no difference between programme conditions for Māori. For this group, the Control had a mean score of 4.80 and the Project K group reported a mean score of 4.64 for PK. Further, the difference between Māori and the contrast group still existed \( (\text{simple slope} = 0.37, t = 4.38, p < .001) \) and, as before, there was no difference within the Project K group.

![Figure 15. Interaction between programme condition and Māori ethnicity for the additional social self-efficacy follow up effect.](image)

A significant Condition.Asian interaction was also found for one year post-programme social self-efficacy but only for the sustained effect \( (b = 0.39, p < .05) \). The pattern of effects was quite different to that found between Māori and the contrast group (see Figure 16). A significant difference between Project K and Control was obtained for the young Asian participants \( (\text{simple slope} = 0.57, t = 3.30, p < .001) \) demonstrating that Project K Asian participants had a higher mean of 4.91 whereas Asian Control students had a mean of 4.35. No difference between Project K and Control was obtained for the contrast group. Here, Control participants reported an average social self-efficacy score of 4.47 while for Project K
it was 4.65. Looking at differences between the Asian and contrast participants within the Control group, there was no significant difference in their scores; however, Asian participants in the Project K group reported higher mean scores than those in the contrast group (simple slope = 0.26, \( t = 1.96, p < .05 \)).

Figure 16. Interaction between programme condition and Asian ethnicity for social self-efficacy one year post-programme.

For the social self-efficacy one year post-programme additional follow up effect only, a marginally significant interaction between Condition and Decile was obtained (\( b = -0.03, p = .06 \)). The simple slopes analyses showed there were in fact no differences between Project K and the Control groups for either low or high decile groups (for low decile: \( M_{PK} = 4.69 \) and \( M_{Control} = 4.75 \); for high decile: \( M_{PK} = 4.76 \) and \( M_{Control} = 5.00 \)). Furthermore, there was no difference between the low and high decile groups within Project K. No \( t \) or \( p \) values could be obtained for the difference between low and high decile participant means within the Control group, which suggests the solution for this interaction effect is not stable. Nevertheless, Figure 17 depicts the predicted means for participants from low and high decile schools for Project K and the Control.
Figure 17. Interaction between programme condition and school decile rating for the additional social self-efficacy follow up effect.

Career Decision Self-Efficacy

The main effect of Condition remained significant in the M2 interaction model (b = 0.33, p < .05) as did the main effect of Asian ethnicity (b = -0.22, p < .05). The only interaction effect obtained for career decision self-efficacy was a marginally significant interaction between Condition and the Asian ethnicity variable (b = 0.35, p = .10). The simple slopes analyses for this interaction demonstrated that the difference between Project K and Control in the contrast group was only marginally significant ($M_{PK} = 3.57$ and $M_{Control} = 3.42$; simple slope = 0.15, $t = 1.69, p = .09$) whereas the difference was significant for Asian participants. Project K participants in this group had higher scores ($M_{PK} = 3.53$ and $M_{Control} = 3.02$; simple slope = 0.51, $t = 2.00, p < .05$). Moreover, a significant difference between Asian participants and participants in the contrast group was obtained within the Control group (simple slope = -0.40, $t = -2.49, p = .05$) with the contrast group scoring higher than Asians but there was no difference between these groups when they had been involved in the programme (see Figure 18).
NCEA Level 1 Credits

When the interactions were accounted for, the main effects of Māori, Pacific and Asian ethnicity remained significant ($b = -17.56, p < .001$ for Māori; $b = -21.97, p < .001$ for Pacific; and $b = 8.81, p < .05$ for Asian). Interestingly, the main effect of Condition and Decile became marginally significant ($b = 9.07, p = .09$ for Condition; $b = 0.86, p = .08$ for Decile) and the main effect of Gender became significant ($b = -5.21, p < .05$).

A significant Condition.Asian interaction was also obtained for NCEA Level 1 credits ($b = 18.63, p < .05$). The simple slopes analyses revealed that no difference existed between Control and Project K for those in the contrast group ($M_{PK} = 74.93$ and $M_{Control} = 75.18$) but Asian participants in the Project K group obtained a greater number of NCEA Level 1 credits on average compared to their Control counterparts ($M_{PK} = 93.06$ and $M_{Control} = 74.67$; simple slope = 18.39, $t = 3.58, p < .001$). In terms of ethnic differences within programme conditions, no difference was found between Asians and contrast participants in the Control group but a difference was obtained within the Project K group (simple slope = 18.13, $t = 2.23, p < .05$). Figure 19 depicts these differences.
Finally, a marginally significant interaction between school decile rating and programme condition was obtained ($b = -1.28, p = .07$) and was further investigated using simple slope analyses. These results showed that for participants in both low and high decile schools, there was no significant difference between Project K and Control participants (For low decile: $M_{PK} = 84.67$ and $M_{Control} = 79.51$; for high decile: $M_{PK} = 85.19$ and $M_{Control} = 88.19$). There was no difference between the decile groups for those participating in Project K and the $t$ and $p$ values could not be generated for the difference between the low and high decile school participants within the Control group. Like the Condition.Decile interaction for the social self-efficacy additional follow up effect, this suggests the estimation for the effect is not stable. Nevertheless, Figure 20 illustrates the predicted means for programme sites with decile ratings falling one standard deviation below and above the decile mean for both Project K and Control.
No significant interactions were obtained for the M2 NCEA Level 1 achievement status model. The main effects of Māori and Pacific ethnicity were still significant (b = -0.78, $p < .01$ for Māori; b = -1.30, $p < .01$ for Pacific). Also, the main effect of Gender and of Decile became marginally significant (b = -0.38, $p = .06$ for Gender; b = 0.06, $p = .06$ for Decile) in the final model with all interactions included.

**OVERVIEW OF KEY FINDINGS IN PLAIN TERMS**

**Programme Effects**

Taken together, the key patterns of findings in the above sections provide strong evidence that, on average, Project K had a consistent positive effect on all of the efficacy outcomes, but no evident influence on the academic achievement outcomes. Looking at the overall rate of change in academic and social self-efficacy from just prior to the programme start date until one year after it was completed, we see that Project K and Control participants began with similar mid-range levels of academic and social self-efficacy. However, the academic self-efficacy of Control participants showed little change over the two year period and their social self-
efficacy increased slightly (less than a quarter of a response category on the 6pt Likert scale per year) over the same duration. In contrast, for the Project K group both outcomes increased at a higher rate (almost three quarters of a response category) than the Control group from programme start to programme finish and for social self-efficacy the increase continued but at a slower rate (about a quarter of a response category). For academic self-efficacy the rate decelerated to a point where it began to descend (by a rate of almost a third of a response category over the next year) yet it did not descend to its original baseline level nor to that of Control participants a year after the programme had finished.

The Project K group also had higher reported levels of academic and social self-efficacy than the Control group when the groups were directly compared at the immediate end of the programme. Project K participants scored, on average, half a response category on the 6 point academic self-efficacy scale higher than Control participants and slightly less than half a response category higher on the social self-efficacy scale.

Corroborating the above statement about the curvilinear trajectory, the differences were indeed still evident one year after programme completion with the Project K group scoring about a third of a response category higher than the Control group on both outcomes. Furthermore, the latter effects were obtained after taking into account participants’ baseline levels of efficacy and these programme effects remained when assessing the differential programme effects for gender, ethnicity, and school decile rating (i.e. when the interaction effects were included).

Likewise, a difference between Project K and Control (in favour of Project K) existed for the one year post-programme academic and social self-efficacy measures when taking into account the levels of academic and social self-efficacy each participant began the process with and the levels each participant reported at the end of the programme. This suggests that the programme had an additional direct influence on these outcomes in the longer term over and above its effects at the immediate end of the programme (about $1/7^{th}$ of a response category).
However these effects were not significant in the final model that also included interactions. This suggests that these programme effects were being accounted for by one of more of the interaction effects, or the additional parameter estimates reduced the power to detect significant effects in this model.

Although it was a cross-sectional outcome, Project K participants also reported higher levels of career-decision self-efficacy one year after the programme (a quarter of a response category on a 5 point scale) had finished and this effect was stable when the interactions were included. However, the groups were more or less equivalent in terms of the number of NCEA Level 1 credits participants obtained and the proportion that attained the NCEA Level 1 certification. Interestingly, there was a marginal difference between the programme conditions when the interaction effects were included. In this case, the predicted number of credits for Project K participants was about 9 more on average than for Control participants.

**Demographic Effects**

In general, females across the sample reported higher levels of social self-efficacy than males both at the end of the programme and one year after, with or without the interaction effects. They also reported higher levels of academic self-efficacy but only one year after the programme had finished and all differences were smaller than a quarter of a response category. In addition, the mean number of NCEA Level 1 credits females obtained was higher than that of males but only when interaction effects were included in the model.

Relative to the contrast group, Māori and Pacific youth obtained fewer NCEA Level-1 credits while Asian participants obtained more. These effects were stable in the full interaction model. For Māori and Pacific youth, these effects were coupled with a lower probability of being awarded an NCEA Level 1 certificate and, for Māori only, with lower levels of post-programme academic self-efficacy (about a quarter of a response category). On the other hand, Māori had higher levels of post and one year post-programme social self-efficacy while Asian youth reported lower levels of career decision self-efficacy (also about a quarter of a
response category). Again the effects were evident with and without including the interaction effects.

School decile rating had an influence (independently of gender, ethnicity, and programme condition) on social self-efficacy at one year post-programme only when investigating the effect over and above the associations with baseline and post-programme levels. On average, participants scored .02 of a response category more on the social self-efficacy scale for every unit increase in school decile rating. This equates to a difference of about a quarter of a response category on the 6 point scale when comparing decile 1 to decile 10 participants. School decile rating also had a marginal influence on both academic achievement outcomes, but only when the interactions were included in the model.

**Interaction Effects**

In terms of the consistency of Project K’s effectiveness across participant subgroups, we saw that it was consistent for the effects on self-efficacy incurred over the duration of the programme, but some subgroup differences did emerge for particular outcomes one year post-programme. Remarkably, Project K seemed to reduce discrepancies that were evident in the overall sample between different subgroups, at least for a few different outcomes. For instance, males, on average, reported lower levels of academic self-efficacy than females; however this effect was not evident in the Project K group. Males in Project K reported higher levels than males in the Control group (about 2/5 of a response category) and there was no gender difference in academic self-efficacy within the Project K group. Similarly, Māori reported higher levels of social self-efficacy than the contrast group one year post-programme but these groups were not different when they were Project K participants, and the Project K contrast group had higher scores than the Control contrast group (about half a response category). Moreover, Asian participants generally reported lower levels of career decision self-efficacy but not if they had taken part in Project K. Asian Project K participants had higher levels of career decision self-efficacy than Control participants (about half a response category) and within the Project K group Asians fared as well as those in the contrast group.
A similar pattern was obtained when looking at the means for participants in low and high decile schools. Discrepancies were more apparent (in favour of participants from high decile schools) within the Control group for NCEA Level 1 credits and social self-efficacy one year post-programme (when baseline and post-programme scores were adjusted for). Unfortunately, the statistical significance of these discrepancies could not be determined. On the other hand, the Asian participants demonstrated the opposite effect for NCEA Level 1 credits. In general, young Asians obtained a higher number of credits relative to the contrast group and this effect appeared to be magnified by Project K. Asians in Project K obtained more credits than Asians in the Control group and more credits than Project K participants in the contrast group. Additionally, Asians reported higher levels one year post-programme social self-efficacy (when only baseline levels were adjusted for) than the contrast participants when they were part of the Project K group and higher levels than their Asian Control counterparts.

In sum, we obtained consistent evidence that: the programme had a positive effect on efficacy outcomes but not on academic achievement outcomes; males tended to have lower efficacy beliefs than females for academic and social domains; and the three most common minority groups found in New Zealand had different strengths and weaknesses when it came to efficacy and academic achievement outcomes across the sample in general. The interaction effects are less consistent and much more complex. That said, the pattern of interaction effects across the different outcomes suggest that the influence of Project K in the long-term (one year after the programme) may function to reduce existing disparities between participant subgroups. Next, I explore the potential reasons underlying all of these effects in conjunction with the limitations, strengths and contributions of the study.

**DISCUSSION**

The primary aim of the current study was to assess the effectiveness of Project K for increasing participants’ academic, social, and career-decision self-
efficacy and for improving their levels of academic achievement as measured by the number of NCEA Level 1 credits participants received and the proportion of participants that were awarded an NCEA Level 1 certificate. Consistent with the first three study predictions, the academic and social self-efficacy scores of Project K participants increased at higher rates than those for Control participants from just before the programme began until one year after it was completed; Project K participants had higher academic and social self-efficacy scores than Control participants at the immediate end of the programme and the difference between the groups was maintained one year after programme completion. Project K participants also had higher levels of career decision self-efficacy at the one year post-programme follow up.

Needless to say, the positive programme effects on self-efficacy outcomes are not surprising. A review of the qualitative results obtained for Project K’s theory of change (Chapter Four) imply that Project K activities are devised in a way that aligns closely with the positive youth development, adventure, service learning, and mentoring programme literature. Furthermore, as outlined in the introduction to this Chapter, the latter three programming types incorporate the four key sources of self-efficacy: personal and vicarious mastery experiences, positive and constructive performance feedback, and the regulation of anxious arousal.

Contrary to the fourth stated prediction, the programme effects for academic achievement outcomes were not as promising as those for self-efficacy. There was no evident difference between Project K and Control in the proportion of participants who had obtained the NCEA Level 1 certificate. Project K participants did, however, report a marginally higher number of NCEA Level 1 credits than Control students when accounting for the interaction effects. The lack of strong evidence for Project K influencing academic outcomes seems incongruent to what would be expected given the ubiquitous finding that academic self-efficacy is positively associated with achievement in academic contexts (e.g. Carroll et al., 2009; Multon et al., 1997; Pajares, 2002; Phillips & Gully, 1997). If Project K boosts academic self-efficacy over the course of the programme and this effect is more or
less sustained one year later, then why did we not see stronger effects on the academic achievement outcomes?

A number of factors could have contributed to this result. First, it is informative to review whether or not academic self-efficacy is indeed linked to the academic achievement outcomes in our sample. A look at the correlation matrix indicates that, at each time point, higher levels of academic self-efficacy were associated with a higher number of NCEA Level 1 credits. Furthermore, higher levels of academic self-efficacy at the end of the programme and one year after were also associated with a greater likelihood of attaining an NCEA Level 1 certificate.

This suggests that Project K may be indirectly linked to improved academic achievement outcomes via its influence on academic self-efficacy. Nevertheless, the absence of a direct association between programme condition and NCEA results implies that much of what occurs within Project K is not specifically promotive of this particular standardised measure of academic achievement. In fact, some aspects of the programme may compete with NCEA progress. For instance, Project K participants are taken out of school for three weeks during the Wilderness Adventure and usually miss some school during the Community Challenge. Project K participants interviewed in a previous study reported having difficulties catching up on missed schoolwork and obtaining the needed support from teachers when returning from the Wilderness and Community components (Warren, 2005). Furthermore, the additional meet ups with mentors and the Project K group over the remainder of the programme may also compete with time that could be spent on homework. Mentors may support Project K participants with homework and academic goals but the degree of academic support is likely to vary substantially across the dyads. Because the accumulation of NCEA Level 1 credits occurred over the duration of the programme, it may have been particularly difficult for Project K participants to surpass the Control participants’ level of achievement on this indicator. The boost in academic self-efficacy that did occur as a result of the programme may be more beneficial to academic success in the longer term.
Increases in social self-efficacy may also generate additional tension between social and academic goals, despite it being important for academic adjustment and retention. Carroll et al. (2009) surmised that the negative association they obtained between social self-efficacy and English academic achievement in an Australian youth sample was linked to the competing interests between academic and social activities at this life stage, and also the fact that high achievers may be viewed as less popular to their peers. So it may be that while there is a shared sense of general efficacy across the academic and social domains, some aspects unique to social self-efficacy may be negatively associated with academic achievement outcomes. It would be informative to more thoroughly investigate the links between Project K and academic achievement outcomes with academic and social self-efficacy as potential mediating variables to illuminate our understanding of the above issues.

Given the inevitability that Project K activities do compete for time with academic responsibilities, something must be said about the absence of iatrogenic effects for the academic (or any) outcomes. One may expect school performance to suffer as a result of the time missed, yet this was not the case. Project K students were able to maintain a similar level of academic achievement with Control students and the fact that their academic self-efficacy received a boost may mean that they have greater academic success in the future. Data have been collected on educational and employment outcomes for this RCT sample three years post-programme. It will be interesting to see if there is greater divergence in these outcomes across the two conditions at this point in time.

Alternatively, Project K’s lack of strong influence on the academic achievement outcomes may simply be the result of insensitivity in the measures used. Stewart and Archbold (1993) explain how the distribution of scores on a measure may reduce the ability to detect changes resulting from an intervention. Including more response options on the measurement scale increases sensitivity to change, but the NCEA Level 1 Status measure employed in this study had a dichotomous response option. Furthermore, the larger the amount of within-group variance relative to the between-group variance, the more difficult it is to detect
real differences between programme and control conditions (Stewart & Archbold, 1992, 1993). In a repeated-measures design, the within-group variance can be adjusted for and this can increase the sensitivity of the measure to a degree (Stewart & Archbold, 1993) but the NCEA Level 1 credits outcome was cross-sectional, credits ranged from 3 to 189 and the standard deviations were large for Project K and for Control (35.10 and 36.33, respectively). Consequently, both measures of academic achievement were less than ideal for detecting group differences.

Magnitude of Programme Effects

While the results for academic achievement are somewhat tenuous, the consistent, significant and positive programme effects for efficacy outcomes put Project K in a favourable light. All the same, Lipsey and Cordray (2000) point out the general acknowledgement within the evaluation research community that significant differences between the programme and control groups do not necessarily translate into practical and meaningful differences. In consequence, reports of significance are generally accompanied with assessment of the magnitude of the effect.

Above, I described the size of the effect in terms of the unstandardized differences between groups on the actual measure used (i.e. the 5 or 6 point Likert scale, number of credits or percentage of certificates attained) but because efficacy is an abstract psychological construct it is still difficult to grasp the practical significance of these results. Generally when using standard regression techniques, the convention is to report the $R^2$ for the model. This provides an estimation of the percentage of variance in the outcome that can be explained by the variables in the model. This can be further broken down into part correlations ($\eta^2$) which signify the amount of variance explained by one variable independently of all others (Field, 2009). This is problematic when using multilevel modelling because there is variance to be explained for multiple variables and at multiple levels. At times, adding a predictor may even generate negative values for the additional variance explained (Snijders & Bosker, 1999); hence why I chose to focus on differences in unstandardized units.
I also provided effect sizes based on the standardised mean difference scores of the raw means for the programme and control groups (accounting for baseline differences when possible). These are merely descriptive and do not account for the hierarchical nature of the data or any demographic covariates. Thus, they are meant to supplement the more precise HLM estimates to allow comparison with similar types of programmes, as this is the standard procedure for calculating the magnitude of programme effects in meta-analytic studies (see DuBois et al., 2002, 2011; Durlak & Weissberg, 2007; Hattie et al., 1997; Lipsey & Cordray, 2002).

The effect sizes associated with Project K’s influence on the self-efficacy outcomes are heartening when compared to those obtained previously for positive youth development programmes, adventure programming, and mentoring. The pre to post-programme academic and social self-efficacy effect size was moderate at approximately .50. This is larger than the mean overall effect of .22 found by Durlak & Weissberg (2007) when they investigated the effects of 73 after-school programmes for young people on personal and social outcomes. Even comparing their results for self-perceptions specifically, Project K’s effects were larger. Durlak and Weissberg obtained an effect size of 0.34 for 22 studies focusing on outcomes such as self-esteem, confidence, and self-efficacy. Furthermore, their finding was larger than those found in previous reviews of similar programmes, according to their comparative assessment (although it should be stated that the programmes in their sample tended to target elementary and middle school aged youth).

The Project K pre-post effects for academic and social self-efficacy were also larger than the overall effect size of .34 found in Hattie et al.’s (1997) meta-analysis of 151 samples of adventure programme participants and the .31 effect size they obtained specifically for self-efficacy outcomes. Likewise, DuBois and his colleagues only obtained an overall pre-post effect size of 0.14 for 59 independent samples of youth mentoring studies conducted between 1970 and 1998 (DuBois et al., 2002) and .21 for 82 samples from youth mentoring evaluation studies conducted between 1999-2010 (DuBois et al, 2011). For both mentoring meta-
analyses, there was no significant variation in the effect size across outcome types
(e.g. educational, social, behavioural).

No meta-analyses specific to service-learning programmes could be
obtained. The reviews that do exist tend not to specify the magnitude of the effects,
though effects are generally in the positive direction (see Billig, 2000, 2004; Conrad

*Follow Up Effects*

The growth trajectory results for academic self-efficacy demonstrated that
the mean for Project K participants was slightly lower at one year post-programme
than immediately after the programme. The difference between Project K and
Control was less but still evident at one year post-programme and, when no
interactions were included, the results suggested that there was an effect of Project
K on one year post-programme academic self-efficacy over and above its influence
immediately after the programme. The pattern for social self-efficacy was similar;
however, the means associated with the growth trajectory effects showed a
continual increase in social self-efficacy from post to one year post-programme for
both groups but there was a slight deceleration for the Project K group. The mean
difference between Project K and Control was, however, larger at one year post-
programme than at the immediate end of the programme when adjusting for
baseline scores, suggesting there was an additional follow up effect.

The *effect sizes* associated with academic and social self-efficacy for the
sustained effect of the programme are slightly less than those immediately after the
programme (in line with the above results for academic self-efficacy but not social)
but they were still significant small-moderate effects (.33 for academic and .44 for
social). The effect size representing the standardised difference in the means for
Project K and Control on career decision self-efficacy was also small to moderate at
.34.

While the additional follow up effect of Project K on one year post-
programme academic and social self-efficacy was not significant for either outcome
when interactions effects were included, the effects were significant when the
interactions were absent. However, the non-significant effect sizes of the programme between the post and one year post measures suggests that the additional follow up effect of Project K obtained using HLM is not practically meaningful. One must of course remember that the effect size is a less robust estimate and that these approximations did not account for baseline levels.

Nevertheless, when comparing Project K’s follow up effect sizes for efficacy outcomes to those found in other related meta-analyses, Project K still appears favourable. DuBois et al.’s (2011) youth mentoring meta-analysis yielded a pre-test to follow-up effect of .17 over 7 studies and a more or less non-existent additional follow up effect of -.03 (when post-test scores were subtracted out of the calculation). Although the follow up length (mean = 23 months post-programme) was generally lengthier than that for Project K. Durlak & Weissberg (2007) found a similar follow up effect of .19 for self-perception outcomes in their meta-analysis of youth after school programmes but it was also only based on few studies (n = 9) and the mean length of time between post-programme and follow up is unclear.

Interestingly, the additional follow up effect size (i.e. post-programme to follow up effect) Hattie et al. (1997) reported in their meta-analysis was a significant though small .17 and for self-efficacy outcomes it was .21. Evidently, this is larger than the additional follow up effects found for Project K on academic and social self-efficacy but at least two caveats apply to this discrepancy. First, the mean follow up period in Hattie et al.’s study was about half the length (5.5 months post-programme) than that for the current study. Thus it may be that adventure programming effects continue to increase between post and one year post-programme but begin to diminish before the subsequent year is up. Second, the cumulative pre-test to follow up effect in Hattie et al.’s study was less for school-based students when compared to adults. The .40 cumulative effect size for their student sample is comparable to the cumulative pre-test to follow up Project K effects on academic (.33) and social self-efficacy (.44). Hattie et al. (1997) also acknowledged that the additive effect of adventure programmes was a unique finding and that, generally, follow up effects resulting from educational
interventions diminish quickly following completion of the intervention or in some cases they may be sustained at the same level.

To sum, when comparing the magnitude of Project K’s influence on efficacy outcomes (using an, albeit, rough estimate of effect size), the programme surpasses what can be expected for pre-post effects of similar programmes. It also seems to fare better in the longer term when compared to mentoring or general after school programmes for youth and, effects are similar to what can be expected from youth adventure programme. Furthermore, the effects obtained (both in terms of HLM estimates and effect sizes) may be conservative estimates as longer-running Project K programmes were less likely to have been included in the sample. Common sense dictates that programmes with longer histories should generally show greater positive effects than newly implemented ones. The comparisons to single-component programmes are impressive; yet we should also keep in mind that Project K combines all three programme types and thus we should expect stronger effects.

**Demographic Main Effects and Interactions with Programme Effects**

Heterogeneity in the effect of a programme across participant subgroups or across different contexts may also wash out the overall magnitude of a programme’s influence (Lipsey & Cordray, 2002; Rhodes, 2008; Rhodes & Lowe, 2009). To elucidate some of these potentially contradictory or reduced effects, interactions between programme condition and demographic factors were explored. Below, I discuss these in conjunction with the overall pattern of effects for gender; Māori, Pacific, and Asian ethnicity; and school decile rating.

**Gender Effects**

When gender differences did exist, the results consistently favoured females; however the effects were not consistent across outcomes. Females did report higher levels of social self-efficacy at the immediate end of the programme and one year after but they only reported higher academic self-efficacy scores at the latter time point. Only in the full interaction model did gender differences exist for NCEA Level 1 credits and, even then, the effect was only marginal.
Gender differences in efficacy beliefs have been routinely examined and the pattern of results in the overseas literature also differs depending on the domain or subject under investigation. Gender differences across studies investigating efficacy for academic subjects and/or for social abilities are rather equivocal (see Bandura et al., 1999; Britner and Pajares, 2006; Carroll et al., 2009; Connolly, 1989; Pajares & Miller, 1994, Pastorelli et al., 2001; Zimmerman and Martinez-Pons, 1990). A more consistent pattern arises with investigations of self-regulation for learning efficacy, with females generally reporting higher levels (Britner & Pajares, 2006; Caprara et al., 2008; Pastorelli et al., 2001). This is, in essence, what the Project K measure of academic self-efficacy reflects; thus, the finding indicates that females tended to feel more confident in their abilities to do things like complete their homework, study for a test, or pay attention during class. This likely also contributed to the marginal gender difference in academic credits.

Little has been done with regards to investigating gender differences in efficacy beliefs within the New Zealand context but research exploring gender differences on other academic outcomes report a similar pattern of favourable effects for females (Ministry of Social Development, 2010; Statistics New Zealand, n.d.; Telford & May, n.d.). In New Zealand, females tend to leave school with higher qualifications (Ministry of Social Development, 2010; Statistics New Zealand, n.d.), have higher reading literacy (Telford & May, n.d.), and report enjoying doing school work more (Adolescent Health Research Group, 2008).

For the current sample, the difference in academic self-efficacy did not emerge until participants were 15 to 17 years old, yet the finding implies that males may need extra encouragement and support to be focused and self-directed in their studies. Interestingly, at one year post-programme, males who were in Project K exhibited similar levels to females in the programme and they reported higher levels than Control participants. Thus, it may be that Project K provides the needed support but the effects are delayed (I return to this later).
**Ethnic Group Differences**

The ethnic group differences in the academic achievement outcomes were consistent with what is generally reported for Māori, Pacific, and Asian students in New Zealand (Boven, Harland & Grace, 2011; Ministry of Social Development, 2010; Statistics New Zealand, n.d.; Telford & May, n.d.). In the current study, Māori and Pacific students obtained fewer credits and were less likely to attain the Level 1 certification whereas Asian participants obtained a higher number of credits. Māori and Pacific youth, generally, are more likely to leave school with no qualifications (Boven et al., 2011; Ministry of Social Development, 2010; Statistics New Zealand, n.d.) and have lower levels of reading, math, and science literacy than individuals identifying as European or Asian (May & Telford, n.d.). Furthermore, in 2008, there was a higher proportion of Asian students leaving school with at least an NCEA Level 2 certificate than any other ethnic group (MSD, 2010).

For Māori, the poorer NCEA results were coupled with lower levels of academic self-efficacy post and one year post-programme. This is consistent with Hollis et al.’s (2011) theoretical argument that Māori exhibit many symptoms of low self-efficacy and these may have resulted from the damaging long-term effects of colonisation. This is also consistent with the theory that efficacy beliefs influence interest, effort, persistence and performance (Bandura, 1977, 1989, 1997, 2006). Thus, the disparity in educational outcomes may stem in part from societal expectations (e.g. racism) that undermine young Māori’s beliefs in their academic (or self-regulation for learning) capabilities. Alternatively, it could be that there is disjuncture between the assessment of academic self-efficacy (i.e. largely in terms of self-regulation) and this particular standardised measure of academic achievement (i.e. NCEA). Perhaps NCEA results are not an appropriate indicator of educational progress for Māori.

It is interesting to note that Pacific youth did not report lower levels of academic self-efficacy despite also obtaining poorer NCEA results. This suggests that the reasons behind the lesser academic performance results for both groups may be different. Given the consistent disadvantage both Māori and Pacific groups
experience with regards to education, employment, and health outcomes, this is something well worth investigating.

Unfortunately, the absence of interaction effects between programme condition and Māori or Pacific ethnicities for the NCEA outcomes indicates that Project K was not effective in changing the relative position of these groups (thus reducing the discrepancies). In fact, Project K appeared to further boost Asian participants’ academic performance because Asian participants within the Project K group did even better in terms of the NCEA Level 1 credits they accumulated than Asian participants in the Control group. They also obtained more credits than other ethnicities within Project K. As I explained earlier, Project K activities may compete with time that could be spent on improving academic abilities. For participants that already struggle within the educational context, it may be especially hard to keep up despite feeling more confident about being able to focus and self-direct their learning. Whereas for those who generally do well (e.g. Asian students), the boost in efficacy that Project K provides may enable them to exceed previous personal levels. Again, it is heartening to see that participation in the programme did not worsen academic achievement outcomes for any group, but the processes underlying these outcomes clearly need to be elucidated.

A converse pattern of results did emerge for Māori’s social self-efficacy levels one year after programme completion, with Māori reporting higher levels than non-Māori. This illustrates a strength of the Māori students in this study in that they had stronger beliefs in their social abilities (e.g. expressing their opinions, making friends and working in teams) than other participants. Given the overwhelming message across New Zealand research reports that Māori are a disadvantaged group (Boven et al., 2011; Clark et al., 2008; Ministry of Social Development, 2010, Statistics New Zealand, n.d.; Telford & May, n.d.), it is important to highlight positive effects (Clark et al., 2008).

The interaction effect between programme condition and Maori ethnicity on this outcome showed that within Project K there was no discrepancy between Māori and those in the contrast yet there was in the Control group. It is hard to tell
what programme mechanism is driving the change, but it could be that, with support from facilitators, Maori participants may pull their PK contrast peers’ confidence in social contexts up towards their own levels.

Another interesting result was the effects associated with Asian participants’ career decision self-efficacy. Their superior academic performance did not translate to confidence in career exploration and decision tasks, as they tended to report lower levels than non-Asian participants. This finding implies that Asian students may have difficulty with the school to work transition, despite doing well academically; therefore, they may require more support with these tasks.

This could be linked to the perception of Asians that they are frequently targets of racism in New Zealand society (Ministry of Social Development, 2010). The anxiety associated with transitioning into the world of work because of this perceived racism may compromise their efficacy in exploring career options and making decisions about these. This fits with Gloria and Hird’s (1999) previously discussed finding that anxiety associated with ethnic prejudice negatively influenced the career decision-making self-efficacy of racial or ethnic minorities in comparison to Caucasian participants. Fortunately, it seems that Project K can provide this support since Asian students within the programme did not suffer from lower career decision self-efficacy relative to the contrast group, while the discrepancy did exist in the Control group.

*School Decile Effects*

The only main effect for school decile rating was on one year post-programme social self-efficacy after adjusting for pre and post-programme scores. The higher the decile rating, the higher were the participants’ scores. This aligns with the general finding that those living in areas of higher economic wealth in New Zealand have better well-being outcomes (Adolescent Health Research Group, 2008; Ministry of Social Development, 2010). It is surprising that we did not find other main effects for decile rating but I address why this may be so in the limitations section below.
Though marginal and inconclusive, results for the interactions between programme condition and school decile rating for post-programme social self-efficacy and NCEA Level 1 credits point to the potential for Project K to reduce disparities between low and high decile students. When part of the Control group, it appeared that high decile students outscored those from low decile contexts but the discrepancy was not evident with the programme group.

**Project K -- a Force for Reducing Disparities?**

The potential for Project K to be differentially effective for participant subgroups based on gender, ethnicity, and school decile rating was also explored to address the secondary study aim. To clarify, this was to gain a better understanding of how Project K fares in terms of supporting the different subgroups within its diverse target population when it comes to improving these outcomes.

Remarkably, the overall pattern of interaction effects implies that Project K may be particularly beneficial for groups that, on average, perform more poorly on the target outcome. The discrepancy between males and females found for academic self-efficacy was reduced in the Project K group, as was the discrepancy between Māori and non-Māori for social self-efficacy, and that between Asians and non-Asians for career decision self-efficacy. The interaction effects between programme condition and school decile rating just described also fit with this pattern.

This finding is further substantiated by results from subsequent analyses of data from this RCT sample. Leeson, Harré, & Deane (2012) separated the baseline sample into groups with low, mid and high total efficacy scores (relative to each other) and assessed each group’s gain in self-efficacy from pre to post-programme. They found that those starting the programme with the lowest levels had the highest gains compared to their Control counterparts. The pattern also aligns with Scales et al.’s (2006) finding that service-learning programmes in the U.S. appear to close the achievement gap between students from high and low socioeconomic schools and the finding from DuBois et al. (2011) that youth mentoring...
programmes are more effective for individuals with some but not extreme levels of risk.

One finding from Hollis et al.’s (2011) study of young Māori experiences within Project K may provide some insight as to why this pattern occurred. As pointed out earlier (in Chapter One and Four), most of the interviewees (5 of 6) indicated that the programme atmosphere was one of ethnic equality in that the expectations were the same for all participants. This allowed them to escape the negative stereotype often associated with being Maori. I explained in the introduction how self-efficacy beliefs are largely influenced by the expectations others have of us (whether signalled explicitly or not) and, when negative, these can undermine our self-beliefs.

By conveying equal expectations to all participants, Project K may shift the negative self-perceptions of some participants towards the higher levels of their peers. However, it appears that the effect is delayed. Perhaps Project K sews the seeds for higher self-expectations but the self-realisations are not made until later down the track. The effect is slightly puzzling and, therefore, should be further explored.

OVERVIEW OF STUDY LIMITATIONS AND CAVEATS

Black Box Evaluations

The above results provide important answers to questions about program effectiveness but one major drawback of standard black box or simple programme-outcome evaluations, as we saw in Chapter Two and Chapter Four, is that they contribute little to understanding why the results were obtained (Astbury & Leeuw, 2010; Chen, 1990; Cook, 2000; Lipsey & Cordray, 2000; Weiss, 1997a, 1998a). Project K’s theory of change described in Chapter Four helps to illuminate some of the possible reasons behind these changes but we are still left in the dark in terms of whether the proposed links between programme processes and outcomes function in the manner proposed. For this reason, subsequent studies have already been implemented to investigate the links between additional
mediating variables identified as important during the logic model building process.

---

The Limitations of Real-World Randomised Controlled Trial Programme Evaluation

When randomised controlled trials of social programmes are implemented within a real-world setting, there are always aspects of the experiment that the evaluator cannot control. Within the programme group, people may get different levels of exposure to the services offered, some people may be more motivated to change than others or other factors may influence how participants engage with or respond to programme services (Lipsey & Cordray, 2000). All of these factors can bias true estimates of the programme’s influence. I have endeavoured to control some of these potential variations within the programme group by only analysing data from participants who received at least 75% of programme services and by explicitly exploring any differential effects based on gender, ethnicity, and school decile rating. There are, of course, many other participant and contextual factors that may have influenced individual success in the programme but other measures were not readily available at the time I came on board for the analysis. Additional moderating variables (e.g. participant engagement and facilitator support) have also been collected with more recent Project K samples to further investigate differential effects (as outlined in Chapter Three).

The interpretation of programme effects may also be distorted by changes in the control group condition. Evaluators cannot prevent control participants from receiving other types of services and if similar services were accessed over the duration of the programme the intended contrast between the control and programme participants may be muted (Durlak & Weissberg, 2007; Lipsey & Cordray, 2000). Furthermore, random assignment may serve to create equivalent groups at baseline (the crux of the RCT’s design) but it is difficult to control participant retention over the course of the evaluation. If the drop-out rates are differential across programme and control conditions (both in terms of numbers and in terms of certain types of participants) the equivalence at later time points may be reduced (Lipsey & Cordray, 2000).
This was an evident problem in the current sample. As the results from Table 4 above and the supplementary analyses in Appendix E indicate, attrition was higher in the Control group at post and one year post-programme (28% and 51% in the Control group vs. 9% and 40% for the Project K group) and there were differences in the pattern of participant characteristics who did not respond to later surveys across the Project K and Control conditions. Control participants who responded to the post-programme survey had higher levels of baseline academic self-efficacy than those who did not whereas Project K participants responding to the one year post-programme survey had marginally higher social self-efficacy baseline scores than those who did not. Across both surveys Māori were less likely to respond when they were part of the Control group and Pacific youth were less likely to respond if they were Project K participants. Control participants identifying with an “Other” ethnicity were more likely to respond at one year post-programme and Asian youth across both Project K and Control were more likely to complete the one year post-programme survey. Europeans were generally more likely to respond at both time points and across both conditions than other ethnic groups. While baseline levels of efficacy and ethnic proportions across conditions were adjusted for when estimating the impact of the programme, the fact that the groups did differ with regards to these characteristics does somewhat reduce confidence in the accuracy and the generalisability of the results. That said, problem-free randomised controlled trials are only likely to exist in the fantasy world of evaluation (Durlak, 1995).

**Ethnicity Results**

The above section clearly illustrates that the results for ethnic differences both in general and in response to the programme should be tempered with a healthy dose of criticism. These results are further weakened in view of the fact that participant numbers were small for Asian and Pacific subgroups to begin with and these diminished further at later data collection time points (e.g. at one year post-programme there were 19 Asians in Project K and 28 in the Control group and 31 Pacific Peoples in Project K vs. 46 in the Control group, see Table 3 in Appendix C). Moreover, the inclusion of all the interaction effects in the final model
substantially increased the number of parameters to be estimated which further reduces the power to detect significant effects. This is likely why the main Asian ethnicity effects and the Condition. Asian interaction effects were not stable across the standard model results and those calculated with robust standard errors (as noted below the relevant results tables). Thus the different pattern of effects for Asian participants may not be a true reflection what could be expected for other groups of Asian participants.

Even the coding of ethnicity is a hotly contested issue in New Zealand research. The total response coding method appears to be preferred because it allows participants to be represented in each ethnic category they identify with and also because it more fairly represents the distribution of these groups within the New Zealand population (Statistics New Zealand, 2005). This is also consistent with how ethnicity is coded in the Youth 2000 and 2007 surveys (Clark et al., 2008; Helu et al., 2009; Rasanathan et al., 2006). However, it is not a true reflection of what exists in the real world as participant frequencies for each category based on this method outnumber actual participant counts. This could be counteracted by using a Single/Combination coding method which includes ethnic combinations but the group sizes would have been even smaller and would have increase the number of ethnicity covariates in the HLM models, both of which would have further reduced statistical power.

Furthermore, the results were interpreted in a manner that suggests that these ethnicities are discrete categories and are stable individual traits that can influence how a group responds to Project K. In fact, ethnicity is a fluid characteristic that may shift over time and across contexts (Statistics New Zealand, 2005) and although the covariance across the different ethnic minority groups was partialled out of the analyses to permit interpretations of the “unique” effects of Māori, Pacific and Asian ethnicities on the outcomes of interest, in reality there is overlap across ethnic categories. Approximately 20% of this sample reported a combined ethnicity but the unique experience of these participants is not included. Great diversity also exists within the Pacific Peoples group (7 different options were combined) and the Asian group (4 different options were combined) but this
is not captured in the above results because having additional groups would have further reduced the power to detect effects. This is acknowledged to be a problem in research investigating Asian (Rasanathan et al., 2006) and Pacific ethnicities (Helu et al., 2009); however, it is difficult to resolve when group numbers are small.

School Decile Rating Results

Similar comments can be made about the decile effects. The data were analysed in a manner that partialled out the shared variance with ethnic variables despite there being an evident over-representation of Maori and Pacific groups in impoverished areas (Clark et al., 2008; Ministry of Social Development, 2005). This allowed me to separate the independent influence of decile from that of the ethnic variables which is important if trying to dissect out the driving force behind the effects; however, in reality they are inherently linked and thus it is not a good portrayal of the more global picture.

Additionally, the results for Decile, both for the main effects and the Decile.Condition interaction effects lacked stability across the standard model estimates and those calculated with robust standard errors. The fact that Sibley’s (2008) utilities for examining the effects of simple slopes could not generate t and p values for some of the group differences generating the Decile.Condition cross-level interaction effects also reduces confidence in the stability of the estimates. The final sample size at this level ranged from 44 to 49 depending on the outcome of interest. This may be sufficient to detect effects when few parameter estimates are required but, as stated previously, power is substantially affected by additional covariates. Had we had a larger number of programme sites to work with, the reliability of the estimates would have been improved.

Problems Estimating Effects for Random Slopes

One of the advantages of using HLM for programme evaluation research is that, if random slopes are included, the results can provide additional insight for programme evaluators beyond answering “Does the programme work, in general?” Allowing the slopes (which represent the effects of different predictors
on the outcomes of interest) to vary across higher level units can illuminate the complexity of what may be occurring “behind the scenes” of the fixed effects (Russell & Sibthorp, 2004; Seltzer, 1994). To illustrate, the main effect of Project K on post-programme academic self-efficacy may be positive overall, but we know little about the consistency of this effect across programme sites. In HLM, we can test whether or not the effect varies significantly across programme sites or if it is generally consistent.

I was able to ascertain whether or not the outcomes varied significantly across programme sites when no predictors or covariates were included; however, the reliability of the estimates for the variance across slopes (programme, demographic, or interaction effects) were too poor to generate meaningful results about the consistency of the effects because power at this level was too low. This information would have been useful to direct future research on potential moderators at this level or to direct qualitative inquiries into programme implementation issues (Seltzer, 1994).

OVERVIEW OF STUDY STRENGTHS AND CONTRIBUTIONS

Contributions to Organisational Learning and Development

The most evident contribution of this study is that it provides evidence of Project K’s effectiveness. Project K’s main objective is to provide young people who are identified as having below average academic and social self-efficacy with experiences that will boost their beliefs in their capabilities to be successful within these domains, and the results demonstrate that this is what the programme does. Findings such as these inevitably contribute to some sense of reassurance for the individuals coordinating and delivering the programmes and of course this should strengthen their case when petitioning for programme funding.

The results should also contribute to the organisation’s learning about where the programme could be further strengthened. It is encouraging to see that over the course of the programme most participants regardless of gender, ethnicity, or school decile context were positively influenced by the programme in terms of the efficacy outcomes. It is even more noteworthy to think that the
programme may help reduce disparities that exist between participant subgroups later on in their lives.

Questions do remain, however, about the programme’s actual impact on academic achievement due to the lack of sensitivity in the indicators used. FYD may well want to investigate other means of evaluating this outcome. At the same time, the results still raise legitimate concerns about the degree to which some aspects of the programme compete with academic responsibilities and/or other educational opportunities. The need for better reintegration of participants into their home and school lives following the Wilderness Adventure and Community Challenge especially was an issue that surfaced a result of Project K’s theory of change findings (Chapter Four) and I return to it here. It may be time to dig deeper into this issue. If the evidence suggesting it is a problem builds, strategies should be devised to better support participants and the key individuals within their social systems with the transition back into their daily lives. This is especially so for participants that already struggle academically, like young Maori and Pacific, perhaps because the social conditions are not conducive to their learning. If such strategies were employed, perhaps we would see a larger contrast in the academic achievement outcomes of Project K and Control participants.

Methodological Rigor and Credibility of Results

Despite the methodological limitations presented above, the study design has many strengths. As discussed in Chapter 1, positive youth development programme evaluations in general, as well as those specific to adventure programming, service-learning, and youth mentoring, are often criticised for lacking methodological rigor (Catalano et al., 2004; Riggs & Greenberg, 2004; Rhodes & Lowe, 2009; Roth & Brooks-Gunn, 2003b) and not investigating moderating influences (DuBois et al., 2011; Riggs & Greenberg, 2004) or long term effects (DuBois et al., 2011; Durlak & Weissberg, 2007; Larson, 2000). The outcome evaluation of Project K’s impact on efficacy and academic achievement outcomes upheld the highest standards for a standard black box evaluation. It employed random assignment of the participant to programme and control conditions; estimates were calculated after adjusting for inequality in the proportions of ethnic
groups and baseline levels of efficacy; the moderating influences of gender, ethnicity, and school decile rating were also investigated; as were longer term effects (at least one year after programme completion). This allowed me to ascertain not only whether the programme was effective overall but also the extent to which it was effective for different groups and whether or not the effects were sustained. The integration of all these design characteristics is a rarity for youth development evaluation research.

The use of hierarchical linear modelling also strengthened the credibility of the results, especially as the existence of significant variation in individual and programme means for several different outcomes empirically justified its use. Analysing these models with standard techniques such as ordinary least squares regression or ANOVA would have likely led to misinterpretations of the results. HLM also allowed me to appropriately model the influence of school decile rating on lower level outcomes and to establish growth trajectories in the two main outcomes of interest.

*Rigorous research studies on youth development programmes are especially important within the New Zealand context because there is a striking paucity of robust evaluation studies in this field. The review of adventure programming in Chapter Three indicated that New Zealand based outcome evaluations of such programmes are characterised either by small samples or focus exclusively on qualitative data. The systematic review of New Zealand mentoring programmes conducted by Farruggia et al. (2010) demonstrated that few quality evaluations of youth mentoring programmes exist, and I was not able to locate even one study which could be categorised as an evaluation of New Zealand-based youth service-learning programme. Yet one of the core principles of the Youth Development Strategy Aotearoa is that “youth development needs good information” (p. 8) and the government relies on well-conducted evaluation studies and general youth development research to fulfil this need (Ministry of Youth Affairs, 2002).*
This is the first known New Zealand-based randomised controlled trial evaluation of a youth development programme and as such the results contribute to this end. While we still have much to learn in the way of which components with Project K generate the positive effects, this study demonstrates that programmes that incorporate similar elements are promising avenues for promoting young New Zealanders beliefs in their abilities. The fact that programmes such as these may also function to reduce discrepancies between social groups in important outcomes is timely information given that a recent report produced by the New Zealand Institute has highlighted the dire need to improve educational and employment outcomes for disadvantaged youth in this country (Boven et al., 2011). It also brings up interesting questions about the participant make-up of groups within youth development programmes. It may seem logical to target resources to the very low end of the socio-economic or risk spectrum (as suggested by Boven et al., 2011); however, if the reduction in discrepancies is a function of other students pulling “lower” students up to a similar level, then this issue deserves careful reflection prior to making related policy or resource-allocation decisions.

Limitations aside, the main effect findings for ethnicity and school decile rating are also thought-provoking. For ethnicity, the findings illustrate that, when it comes to self-perceptions of competency, the main ethnic groups found in New Zealand appear to have different needs and strengths. These self-perceptions also relate differently across efficacy domains and with academic achievement outcomes. This kind of information is becoming even more pertinent as the proportion of Asian (Ministry of Social Development, 2010; Rasanathan et al., 2006;) and Pacific peoples (Helu et al., 2009; Ministry of Social Development, 2005) in New Zealand society continues to grow and the youth population especially is growing more diverse (Adolescent Health Research Group, 2008; Dunphy et al., 2008). More research is needed to clarify the factors underlying these differences, therein lies the key to reducing social disparities.
CONCLUSION

In general, much remains to be done with regards to elucidating the mechanisms behind individual growth in Project K and other youth development programmes as well as the additional participant or contextual factors that may interact with programme processes. Some of this work is currently underway – both through further investigation of this RCT sample and on new Project K programme samples. Nevertheless, this study produced a rigorous account of Project K’s effectiveness on efficacy outcomes and it highlighted interesting areas for further youth development research, especially with regards to how PYD programme processes may interact with participant characteristics to reduce social disparities.
Chapter Six

SELF-REFLECTIVE META-EVALUATION

THoughtful Evaluation IN ACTION

…it is also appropriate for an evaluator to plan and carry through even small-scale formative evaluations with a metaevaluation mind-set.

~ Daniel Stufflebeam (2001b, p. 204)

THE META-EVALUATIVE MIND-SET

Evaluations, like the programmes they investigate, are subject to fallibility and thus should also be assessed against evaluative standards, a concept known as meta-evaluation (D. L. Stufflebeam, 2001). However, independent meta-evaluations are not always possible. D. L. Stufflebeam (2001) insists that, at the very least, evaluators should self-evaluate. It is with this imperative in mind that I write this chapter.

In Chapter Two I discussed the evaluation standards I established for myself before I began this research journey. These were encompassed by the credible and empowering dimensions of the Thoughtful Evaluation model. With a Thoughtful approach, a single evaluator takes responsibility for holding the balance between the credible and empowering aspects of the evaluation. The premise of the thoughtful evaluator incorporates, at its core, the idea that self-reflection should occur throughout the evaluation process. This allows the evaluator to make adjustments if the process becomes too focused on the aspects of one dimension to the detriment of the other. Below I draw attention to how the credible and empowering aspects were attended to, I discuss the advantageous of this approach in terms of use and then explain the difficulties I encountered in trying to follow the Thoughtful Evaluation model.
CREDIBLE FEATURES OF THE CURRENT RESEARCH

To reiterate what I established in Chapter Two, the credibility dimension of the *Thoughtful Evaluation* model includes features generally associated with conducting research with high scientific integrity. More specifically, this suggests that evaluations should be approached in a systematic and technically adequate manner such that the results and interpretations are accurate and defensible. The scientific credibility aspects of the RCT outcome evaluation study depicted in Chapter Five are clear-cut. The RCT design is the strongest methodological option for drawing conclusions about programme effectiveness (Campbell & Stanley, 1963; Cook, 2000; Cook & Campbell, 1979; Lipsey & Cordray, 2000; Shadish, et al., 2002; Trochim & Donnelly, 2008; Weiss, 1998a). The self-efficacy outcome measures were determined to have sound psychometric properties (and thus construct validity and reliability) prior to inclusion in the evaluation (Moore, 2005) and objective measures of academic achievement (i.e. NCEA results sourced directly from schools or the Ministry of Education) were included to add additional credibility to the design.

Although I was not involved in the initial evaluation design decisions, I endeavoured to further strengthen the credibility of this study in a number of ways. For instance, I re-assessed the psychometric properties of all self-efficacy subscales to ensure the construct validity and reliability held up with the sample of interest and across measurement time points. Robust analytic procedures were purposefully selected to generate more precise estimates of programme effects. In addition, the limitations of the design, measures, and analysis were carefully considered and incorporated in interpretations to provide an accurate account of what the data represent.

My approach to the logic model study presented in Chapter Four also incorporated several strategies that are recognised by experts as contributing to scientifically credible qualitative research (Braun & Clarke, 2006; Patton, 2002; Weiss, 1998a). I analysed the data in a systematic manner following the guidelines for thematic analysis outlined by Braun & Clarke (2006). I used triangulation to assess the degree of convergence across multiple perspectives, including the views
of different individuals within and across stakeholder groups and the consistency of findings across different types of data sources (e.g. interviews, open-ended survey responses, documents). I also cross-checked the findings with stakeholders (including some that took part in the study and others that did not) to verify the validity of the interpretations and to inquire about any important omissions.

The reported findings for each study are presented in a balanced way in that programme strengths are highlighted while also making clear the opportunities for programme improvement. The processes for each study are detailed in full and thus open to critique from others. Finally, the findings contribute to a wider theoretical knowledge base (e.g. positive youth development and programme evaluation), which is another feature promoted within the credibility dimension (more on this in Chapter Seven).

**EMPOWERING FEATURES OF THE CURRENT RESEARCH**

Empowering processes are those which increase critical awareness of the operational environment, engender feelings of agency in terms of dealing with this environment, and provide opportunities to enact those beliefs so progress can be made towards the ideal (Zimmerman, 2000). Within an evaluation context, the above processes are facilitated if a wide-range of stakeholder values and views are attended to, knowledge is shared in a respectful and bidirectional manner, decisions are made collaboratively, the evaluation is relevant to the varied groups of stakeholders, and opportunities for participation toward shared goals are provided.

Table 23 highlights how the specific processes detailed in the Chapter Two chronology of the evaluation process can be categorised into four broad empowering strategies. The first, *consultation and collaboration*, reflects the importance of genuine consideration of stakeholder values and needs in combination with collaboration on important evaluation decisions and activities. Framing programme information gathering as *soliciting stakeholder expertise* demonstrates respect for stakeholders by involving them meaningfully in the co-construction of programme knowledge. Providing *open access to evaluation*
information and disseminating findings in an understandable and engaging format are essential for helping stakeholders to develop a critical awareness of their operational environment. Finally, partnering in programme development is about sharing knowledge and collaborating towards actual positive change for the programme and the organisation. Whereas the scientific credibility aspects are largely incorporated in the methods sections of the two core thesis studies, the empowerment aspects are perhaps harder to evidence; thus the processes outlined in Table 2 are accompanied by example outputs that are intended to serve as an empowering process audit trail.

Table 23. Empowering Evaluation Processes

<table>
<thead>
<tr>
<th>Empowering Strategy</th>
<th>Example Processes</th>
<th>Example Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consultation &amp; Collaboration</td>
<td>Continual one to one consultation and collaboration with FYD Executive Trustee &amp; Evaluation Coordinator</td>
<td>Discussion paper reviewing evaluation process with recommendations for changes</td>
</tr>
<tr>
<td></td>
<td>Group consultations with FYD Research and Evaluation Unit members</td>
<td>Revised evaluation measures and process for 2010 and 2011 Evaluation manuals and document to Community Partners detailing rationale for changes</td>
</tr>
<tr>
<td></td>
<td>Consultations with Project K Programme Directors and staff from across the regions</td>
<td>Evaluation questions voting survey and document outlining proposed research plan based on voting results (see Voting survey in Appendix F)</td>
</tr>
<tr>
<td></td>
<td>Stakeholder voting process to determine evaluation focus</td>
<td></td>
</tr>
<tr>
<td>Soliciting Stakeholder Expertise</td>
<td>Logic model workshop with FYD National Support Office staff</td>
<td>Summary of logic model</td>
</tr>
<tr>
<td></td>
<td>Logic model focus groups with Project K staff</td>
<td>Document soliciting feedback from FYD staff and programme practitioners on each component of the logic model</td>
</tr>
<tr>
<td></td>
<td>Invitation to submit feedback on findings/provide own interpretations</td>
<td></td>
</tr>
<tr>
<td>Dissemination &amp; Open Access to Information</td>
<td>Establishment of FYD and Programme Practitioner Communication Network</td>
<td>Written reports with graphs for regional pre-post-programme preliminary RCT results</td>
</tr>
<tr>
<td></td>
<td>Written Reports (with visual aids)</td>
<td>Written report with graphs for National programme pre-post</td>
</tr>
</tbody>
</table>
### Empowering Strategy | Example Processes | Example Outputs
--- | --- | ---
Oral Presentations (with visual aids) to FYD and to programme practitioners | programme preliminary results
Oral presentation (with PowerPoint) to FYD staff on National programme preliminary pre-post programme results
Written report and Programme Advisory Group presentation summarising key findings from preliminary RCT and qualitative results
Written report & oral presentation (with PowerPoint) to stakeholders outlining preliminary Programme Logic Model | Oral presentation of key RCT results and finalised Project K Logic model findings to FYD staff and programme practitioners
Written report & oral presentation (with PowerPoint) to stakeholders summarising key findings from preliminary RCT and qualitative results | Report summarising stakeholder feedback on logic model results

**Partnering in Programme Development**

Programme Development Opportunities Workshop Activity

Table summarising FYD staff and Project K practitioner feedback on programme development opportunities (see Table 1 in Appendix F)

### SUCCESSES OF THE APPROACH

The intended corollary of employing a *Thoughtful* approach is increased utility of evaluation information for internal as well as external evaluation stakeholders. The utility to external stakeholders (i.e. policymakers, practitioners of other programmes, or other researchers) is strongly linked to the perceived credibility of the research and its dissemination through relevant outlets. In time, the peer-review publication process will determine to a large extent the degree of utility my thesis research has had for these external groups; thus, here I focus on the benefits I observed for internal stakeholders (namely FYD and Project K staff).

Process use arises from stakeholder involvement in the evaluation process (Patton, 1997). Patton posits that any substantial stakeholder involvement in the evaluation process inevitably ends up being limited to a small group of individuals who have demonstrated a personal interest in the process and findings. He called
these stakeholders the “primary intended users” (Patton, 1997, p. 41). It became apparent early in the research process that the primary intended users of this Project K evaluation were a few key individuals from FYD’s staff, including the Executive Trustee, the Evaluation Coordinator, the Programme Development Managers, and individuals involved in managing the FYD database. Although, it was not the wide-ranging group involvement I had originally envisioned (more on this when discussing the difficulties below), it was encouraging to see evident process use within this group of intended users. Variations of the logic model development process have been employed by FYD staff to clarify the theory underlying other FYD programmes. In addition, one of the intended users expressed to me that, as a result of her experiences with this Project K research project, she now regularly consults with programme practitioners and encourages their feedback prior to making evaluative decisions.

I have also observed evidence of conceptual use. For instance, recent feedback from one of the newer programme practitioners indicated that the explanation of the programme logic model helped to clarify how Project K was meant to operate in terms of core processes and the key influences that should be attended to.

With regards to instrumental use, the findings from the logic model study have already been used to direct other evaluation activities (i.e. the Project K process-focused follow up study mentioned in Chapters Two and Four). The efficient use of these findings for evaluative direction can largely be attributed to the positive working relationship that was developed with the FYD intended users over the course of this project. The programme development opportunities workshop (listed in Table 23 above) will hopefully also generate more dialogue and eventual instrumental action to modify relevant aspects of programme policies and practice.

It is important to note that the benefits of the process were bidirectional. I gained a wealth of evaluation practice knowledge from the experience. This was
greatly enhanced by the collaborative nature of the research because I could properly appreciate the realities of FYD’s and Project K’s programming worlds.

**INTERESTING TENSIONS**

The *Thoughtful Evaluation* approach *in action* was far from the participatory process I had originally envisioned. Some of the difficulties I encountered were the result of expected tensions between some facets associated with credibility and others linked to empowerment. Other impediments were associated with the characteristics of my particular situation and the evaluation context, although most of these issues were connected. I begin with the former.

**Dimensional Tensions**

When discussing some of the intrinsic tensions between the scientific credibility and stakeholder empowering features of the *Thoughtful* approach in Chapter Two I mentioned that rigorous designs (such as RCTs) can present particular difficulties because of the time they take to generate practical findings and the intensive and intrusive nature of the data collection process. These factors can engender feelings of disempowerment for stakeholders and reduce their levels of cooperation with the process (Weiss, 1998a). Additionally, lengthy surveys distributed frequently throughout the programme can be frustrating for the programme participants and can detract from the positive experience of the programme. These issues can compromise the quality of the data obtained and, consequently, the credibility of the findings.

All of the above concerns surfaced during Project K’s RCT evaluation study. Accordingly, there was a need for consultation with programme practitioners across the regions to build trusting relationships before initiating any new projects. The consultation process resulted in changes to FYD’s evaluation policies (as indicated in Chapter Two) which required making some compromises on the extensiveness of the data collection activities and modifying the roles of the practitioners in the evaluation process to assuage the burden on their other responsibilities. This was a clear lesson that mindful awareness of the inherent tensions between certain dimensional features on the part of the evaluator and
making small compromises on some features in view of others is necessary to keep the Thoughtful Evaluation balance in check.

Other Situational and Contextual Impediments

Stakeholder Involvement and the Empowerment Paradox

As indicated in Chapter Two and above, my initial intention for this research was to develop a highly participatory evaluation of Project K. However, the lack of support received from Project K practitioners when I presented this intention severely limited my ability to conduct an evaluation of this nature. This revealed an interesting paradox for my empowering approach: Empowerment is contingent on active and meaningful involvement on the part of stakeholders yet forcing participation from uninterested stakeholders goes against all notions of empowerment.

Evidently, this is not an uncommon scenario. Smith (1993) and Weiss (1998a) have described scenarios where, even when given the opportunity, stakeholders had no interest in being involved in evaluation activities and this did not appear to influence the credibility and utility of the evaluation. Perhaps the simple gesture of an invitation and opportunity to participate conveys to stakeholders that their values and opinions are respected. In this way, this alone can be empowering.

This realisation on my part meant I had to go back to the drawing board to reflect on other possible avenues to ensure the findings would be credible and useful to internal stakeholders without their extensive involvement because, ironically, this would have been disempowering. In the end, it was more feasible and profitable to focus the collaborative aspects on the primary intended users while involving programme practitioners only more peripherally.

Following the first series of consultation visits to the programme regions, I established a Communication Network which acted as an open communication channel to keep regional stakeholders abreast of developments in the Project K evaluation and to provide access to me so stakeholders could express their
opinions, ask questions and voice concerns. I also sought their opinions of how to focus the evaluation on their needs using an evaluation questions voting survey (see Chapter Two and Table 23 above). When their involvement was re-framed as voluntary and peripheral, the programme practitioners generally became cooperative and forthcoming in providing detailed information about their programme experiences (with the exception of individuals from one focus group). I believe this change in receptivity was partly due to my willingness to modify my approach to accommodate their needs. Again, this demonstrates that the evaluator’s capacity to be reflective and flexible is at the heart of the Thoughtful approach. This is the key to maintaining the credibility-empowerment balance.

The Structure and Nature of the Organisation

Some of the difficulties in employing the Thoughtful approach can also be attributed to the organisational structure. The geographical dispersion of stakeholders across 11 different programme regions made it very costly, both in time and resources, to involve many stakeholder groups (even peripherally) in a personalised and face to face manner. As a result, I was only able to travel to the regions on two occasions and this made it more difficult to build trusting and collaborative relationships.

The hierarchical nature of the organisation also created minor barriers for timely dissemination of evaluation information through the Communication Network because the communication channels were controlled to a large extent by what could be considered to be “gatekeepers”. Early negotiations with the FYD primary intended user group stipulated that information about the programme needed to be reviewed by at least of one person from this group prior to wider dissemination. This generally worked to benefit other stakeholders as the findings were first reviewed by individuals who had much more knowledge about the programme and thus were in better positions than I to evaluate the ways that downstream stakeholders might react to the information. Their input was invaluable in terms of mitigating unnecessary negative consequences due to the way the information was presented and this input also helped ensure the information was succinct and understandable. Nevertheless, it did reduce the
timeliness of communications to these groups and the process did not operate according to the equal and open access communication channel first envisioned.

At the beginning of the process I anticipated involving representatives from all key programme stakeholder groups in the evaluation process (i.e. FYD and Project K trustees, FYD and Project K staff, Project K participants, their caregivers, and mentors, as well as relevant school personnel). However, at another organisational level, the programme practitioners were “gatekeepers” of the communication channels to programme stakeholders further downstream, such as the Project K participants, their caregivers, and Project K mentors. I would have liked to include these individuals but did not have access to their contact details due to ethical agreements between FYD and these individuals and, unfortunately, access to these groups would have meant greater involvement from programme practitioners which they had vocalised clearly that they were against. Instead, I had to rely on indirect contact with these groups via the programme practitioners and this was of course dependent on how much time they had to do this. A clear example of this was when I encouraged the distribution of the evaluation voting survey to other stakeholder groups and programme practitioners from only one programme region passed the survey on to Project K participant caregivers and mentors. As a result, the votes were not a fair representation of stakeholder views from all regions. Additionally, the voices of the young Project K participants in the creation of the logic model could only be included via comments they had previously made on an FYD evaluation survey. This greatly limited the richness of what is an incredibly important programme perspective.

There was also a fairly high rate of staff turnover within the regional Community Partners over the time I was involved in this project. This not uncommon for the not-for-profit sector and when I came on board several of the programme directors had been in their positions for a number of years; thus it is not entirely surprising that many had moved on to different opportunities by the time the project was wrapping up. This, in combination with the lengthy time it took to produce the finalised results (due in part to my lack of methodological skills which are detailed below), meant that the stakeholders who had joined the
network at the beginning of the process were no longer part of the organisation upon project completion. Staff turnover also presented difficulties for verifying the qualitative statements produced by particular individuals and meant that the information needs of the individuals who voted on the evaluation questions survey were perhaps different to those who were currently involved.

**Evaluator Characteristics**

I was a novice to programme evaluation when I began this process and, to an extent, my inexperience hindered my ability to produce a higher quality evaluation. In particular, my lack of expertise in diverse research methods required me to spend a great deal of time building my competency in the appropriate techniques. This reduced the efficiency of what was an already going to be a time-consuming evaluation process. Chen (1994) explains that it is not unusual for evaluators to base evaluation designs on the methods they feel the most comfortable with but he chastises those who take a methods-driven approach rather than choosing methods that best suit the purposes of the evaluation. Rather than employing less suitable or less precise methods, I opted to take more time in order to produce more scientifically credible findings. To compensate, I did disseminate preliminary findings using less adequate methods throughout the process (always with a caveat that these were only preliminary). Shadish et al. (1991) comment that evaluators following contingency-based third stage approaches often choose to report lower quality findings in order to satisfy stakeholders’ need for timely and relevant information.

I also lacked a genuine understanding of New Zealand’s socio-political, historical and cultural context when I embarked on this research journey. I did not become aware until quite late in the research process of the important implications the Treaty of Waitangi, in particular, has for evaluation practice or for youth programming in Aotearoa New Zealand. I did explore the differential effectiveness of the programme for different ethnic groups and I sought consult on the way the information pertaining to Māori was presented to increase the likelihood that this was framed in a manner that accounts for a Māori view. However, this was done retrospectively and I did not purposefully seek to include
the unique view of Māori stakeholders in the qualitative study, other than the incorporation of Hollis et al.’s (2011) findings from their interviews with Māori Project K graduates (a project on which I was a collaborator).

The fact that the Aotearoa New Zealand Evaluation Association’s (ANZEA; 2011) Evaluator Competencies framework was not published until late last year evidently contributed to this weakness in my approach. Interestingly, the disconnect between encouraging cultural consideration at a policy level and actual cultural competency in practice was also observed within the FYD organisation (see Chapter Four). Had cultural consideration been embedded in the organisational culture then I would have likely been made aware of this at an earlier stage. The systematic review of mentoring in New Zealand (Farrugia et al., 2010) suggests that this may also be the case for wider youth development practice in New Zealand. Like the ANZEA Evaluator Competency framework, the Youth Development Strategy Aotearoa (YDSA) is undergirded by the Treaty principles and it was launched 10 years ago; however, the rate of information transference from policy to practice is apparently a slow-going process.

**CONCLUSION**

The *Thoughtful Evaluation* approach was by no means easy to employ in practice. I found the scientific credibility aspects easier to uphold in some respects because these were generally aspects I could control independently. In contrast, the empowering dimension is inherently tied to the quality of the relationships formed with stakeholders and managing the relational aspects of the evaluation process was especially challenging because attitudes were slow to change and relationship quality varied greatly depending on shared values, personality characteristics, and previous personal experiences with evaluation. However, the effort it took to attend to this dimension was worthwhile. There is evidence that the process and findings were useful in several different ways. Furthermore, despite there being natural tension between some aspects of each dimension, attending to the empowering aspects largely enhanced the credibility of the findings because of the shared space in the model.
Chapter Seven

FINAL THOUGHTS
A RESEARCH REPRISE

Youth development needs good information ~ The Youth Development Strategy Aotearoa (Ministry of Youth Affairs, 2002, p. 8)

The primary intention for this research was to produce credible knowledge about Project K in a manner that empowered programme stakeholders to address their information needs. Another intention was to produce this knowledge in a manner that contributes to two threads of “cumulative wisdom” (Donaldson, 2003, p. 134): positive youth development (PYD) programming and programme evaluation.

Consultation efforts early in the research process determined that the Project K stakeholders’ most pertinent information needs centred on knowledge about programme processes (e.g. How does the programme work?), influencing factors of programme success (e.g. What factors facilitate or impede programme success?), and programme effectiveness including differential effectiveness for participant subgroups (Does Project K meet its primary objectives of increasing self-efficacy and for which types of participants is it most effective?). I determined that developing Project K’s theory of change was the most appropriate place to start because this type of theory-driven approach clarifies the important programme elements and influencing factors in a way that incorporates the expertise of those closest to the programme. A theory-driven approach can also demonstrate whether there is agreement in the theory across the organisation and if the theory aligns with the relevant social science literature. In Project K’s case, much evaluation data had already been collected but it had not yet been integrated as part of a coherent programme theory. Thus, it also made sense to explore this data to see how it could contribute to the programme theory and understanding the actual process of change.
Taken together, what resulted was a comprehensive mixed-method evaluation of Project K. I initiated the theory of change development process by triangulating multiple qualitative sources of information about Project K to determine the key theoretical mechanisms and moderators of change (Chapter Four). Additionally, using randomised controlled trial (RCT) data collected by FYD and Project K staff, I employed rigorous statistical techniques to produce a robust account of Project K’s impact on self-efficacy and academic achievement outcomes (Chapter Five). While the programme theory validation process (refer to Figure 3 in Chapter Four) is not complete, we (the evaluation research team made up of individuals from the Foundation for Youth Development and the University of Auckland) have already initiated an investigation to test other hypothesised links.

The combination of a theory-driven approach with an experimental design (quasi or randomised), as seen here, represents an advanced level of evaluation design integration. As stated in Chapter Four, Cook (2000) encourages this kind of integration as it is useful for answering both process and outcomes-focused questions and I propose that this be the “new gold standard” to which evaluators aspire. In developing a novel approach based on the lessons of the programme evaluation field’s past to guide my practice and also explicitly outlining the evaluation process (including why certain decisions were made, the difficulties encountered, and the benefits obtained) I also contribute to reducing the widely-acknowledged gap between evaluation theory and practice (see Chen, 1994; Donaldson, 2003, 2007; Mark, 2003; Rogers, 2007; Shadish, 1994; Shadish et al., 1991; Smith, 1993; Weiss, 1997a, 1997b).

In this final chapter, I integrate the most interesting insights from the programme theory development process, the randomised controlled trial process, and the implementation of the Thoughtful Evaluation approach to highlight the key findings and lessons learned. I begin by illustrating how what we (the programme stakeholders and I) learned about Project K can advance the field of youth development programming – both overseas and in New Zealand. I then reiterate what this thesis offers to other evaluators. Finally, I contemplate specific recommendations for FYD and Project K’s learning and development.
ADVANCING KNOWLEDGE BEYOND THE LOCAL SCENE

Taking the Youth Development Programming Field Forward

The literature review in Chapter Three highlighted the need for more rigorous youth development programming research. Some youth programming researchers have bemoaned the fact that few evaluations include an adequate comparison group or include follow up assessments (Catalano et al., 2004; DuBois et al., 2011; Riggs & Greenberg, 2004; Roth & Brooks-Gunn, 2003b). Assessments of programme processes and moderating effects are also not as common while standard black box pre to post-test impact assessments are ubiquitous (Larson, 2000; Riggs & Greenberg, 2004; Roth et al., 1998). The lack of methodological rigor and process-focused evaluation is especially so in the New Zealand PYD research sector.

By addressing each of the above methodological concerns as well as providing thorough reviews of the youth development programming field, this thesis suggests how programme processes, influencing factors and outcomes are linked. Overall, the combined results from both studies suggest that Project K is a top contender for “model” programme status, both in terms of best practice principles and programme effects. Thus youth development programme designers, practitioners, policymakers and researchers could look to these strengths to enhance the functioning of other youth programmes. The studies discussed here also present intriguing opportunities for further research; I point these out within the relevant sections. Then in a separate segment I further detail the research opportunities.

Project K and Evidence-Based Theory

From its very inception Project K was based on what was known about best practices for youth development (Dingle, 2005; Dingle & Wilkinson, 2011). Since then, FYD has attempted to stay abreast of research developments in the field so that programme development is continually informed by evidence-based theory. In Chapter Five I explained that self-efficacy beliefs are a powerful motivating force because they determine to a large degree what we choose to pursue, how
much effort we expend and how long we persist in striving towards goals (Bandura, 1977, 1989, 1997, 2006). Fostering efficacy beliefs during adolescence is especially important because the many challenges young people face can undermine their beliefs in their abilities (Bandura, 2006). The fact that Project K aims to provide efficacy-enhancing supports to young people with below average academic and social self-efficacy fits well with this need in adolescence and with the competency-building theme that is central to the positive youth development (PYD) perspective.

The theory of change study illustrated that Project K provides structured opportunities for experiential learning and scaffolded support which enables young people to achieve their goals. The processes emphasised by programme practitioners aligned with nearly every best practice principle of PYD, adventure, service-learning, and mentoring programmes. Furthermore, these processes also fit well with self-efficacy theory in that they incorporate the four main sources of efficacy beliefs (experiences of mastery, vicarious mastery experiences, verbal feedback, and anxious arousal). However, Project K takes PYD programming to the next level: By integrating the best practice principles of each programme type and packaging them into one programme using an iterative and cyclical process it maximises the likelihood that participants receive a sufficient dosage of support from multiple sources and that they progress towards greater self-determination. I called this concept “ongoing input”. Other youth development programmes could explore whether ongoing input could be incorporated within their structure to yield greater benefits. Given its novelty, however, the role of ongoing input should be researched more extensively.

The importance of participant reintegration upon completion of youth development components/programmes is another interesting insight revealed by this research. Participant reintegration was identified as an opportunity for Project K’s future development (recommendations are discussed in more detail further on) but is rarely addressed in the youth development literature. Participant reintegration is especially important for services like adventure programmes where participants have intense experiences away from their normative
environment but is also relevant to other types of programmes that have a concrete end. This issue is briefly touched on in the mentoring literature where attention is driven by the need for proper relationship dissolution (see Keller, 2005; MENTOR, 2009); however, I encourage more research on this important issue.

Additionally, in targeting a mid-risk group of young people, Project K stakeholders argue that they are unique in addressing a gap in New Zealand youth service provision. It was very interesting to note that recent research on youth mentoring has shown that those falling more towards the middle of the risk continuum are the individuals that gain the most benefit from mentoring (DuBois, et al., 2011; Schwartz et al., 2010). Thus in identifying this gap Project K appears to have focused on the sector that will most benefit from this this type of support. A better understanding of whether mid-risk youth benefit more than others in adventure and service-learning programmes would of course strengthen the notion that Project K is targeting the most appropriate sector of young people; thus future adventure and service-learning programme research could look into the impact on participants with differing levels of risk factors.

**Programme Effects**

The results from the randomised controlled trial outcome study confirmed that the programme is effective in promoting efficacy beliefs across academic, social, and career development domains. The magnitude of the programme effects for these outcomes are comparable to those obtained for youth adventure programmes and are larger than those generally found for mentoring or general after-school programmes (see DuBois et al., 2002, 2011; Durlak & Weissberg, 2007; Hattie et al., 1997).

The qualitative evidence from Project K participants obtained further substantiated that Project K improves self-concept, especially the competency-related aspects; 63% of the youth respondents reported self-concept gains and for 51% of the full sample these were linked to self-confidence. Moreover, the logic model results pointed to eight other important outcomes that the programme was perceived to deliver (knowledge/skill development, motivation and achievement,
Final Thoughts

connectedness, positive outlook, new experiences, affect-related outcomes, physical health, and maturity). Altogether an impressive 99% of the youth participants reported some kind of gain. These outcome results suggest that PYD programmes incorporating similar principles may obtain similar benefits.

The finding that Project K appears to have a disparity-reducing effect between some participant subgroups was an intriguing revelation. Notably, in addition to targeting the mid-risk gap in youth services, Project K has always been about bringing together a mixed group of young people. Perhaps the active ingredient of the disparity-reducing effect is the “pulling-up” influence that some individuals can have on others, especially when supported by facilitators in a positive and structured social environment. Evidently, the mechanisms underlying these interesting effects need to be further researched but this finding intimates that targeting services to a limited bracket (i.e. only those with extreme needs) within the youth population may not generate the same benefits.

By investigating participant characteristics like gender, ethnicity and socioeconomic location these studies reinforced that different participant subgroups have different levels of efficacy and academic achievement. Disparities in academic achievement for young Māori and Pacific youth relative to other ethnic groups have been documented but less is known about effects related to efficacy beliefs. Additionally, by including school decile rating as a covariate in the RCT analyses, the results revealed that the concerns for Māori and Pacific youth exist separately from those associated with the socioeconomic standing of the schools they attend. These findings signal a need to explore other factors within the social context such as institutionally, personally-mediated, and internalised racism as described by Jones (2000, 2001). For Asian young people it is possible that similar racism-linked factors, while not evident in the educational context, could be impeding the development of a positive vocational identity and career path. There is clearly a need for more research on the factors contributing to ethnic disparities in well-being outcomes for young people to begin with, especially for the growing but understudied Asian and Pacific populations.
Related to this, we learned from the logic model study that cultural consideration is an important element for enhancing the positive experiences of young people in the programme. However, there is evidently a need for more workforce training on cultural competency within FYD and this appears to be a common issue for other programme designers and practitioners in New Zealand (see Farrugia et al., 2010). As previously noted, this aspect was lacking in my approach to evaluation because I did not have appropriate training. There is a disjuncture between the advocacy of cultural competency at a policy level and knowing how to incorporate this into practice. Workforce training in cultural consideration for both youth development and evaluation practitioners is needed. The very recent press release from the New Zealand Minister of Health promoting the first online cultural competency training course for health workers (Health Workforce New Zealand, 2012) highlights the contemporary nature of this important issue. While New Zealand has a unique cultural mix, it is likely that these issues will resonate with evaluators in other multi-cultural settings.

Opportunities for Future Research

Despite my efforts to work closely with stakeholders throughout this research and penetrate the “black box” that is Project K, there is still much to discover about how the programmes processes, influencing factors, and youth development outcomes are linked. Direct observations of Project K in action would be particularly valuable. Donaldson (2003, 2007) also promoted programme observation as a means of discerning whether stakeholder-derived programme theories matched actual programme delivery. Additional research on the effects for participants who did not obtain the full degree of programme services would also be extremely valuable. Their voices are absent from this story and could reveal important insights about any unintended negative effects.

Taking Evaluation Theory and Practice Forward

In 1994, Shadish remarked that when evaluators obtain employment they, more often than not, get straight to the business of doing evaluation with little thought to the theoretical underpinnings of their approach. Particular evaluation practice decisions can advance different values around what information is
“truthful” and “useful” (as I pointed out in Chapter Two). Practicing evaluation with integrity requires self-reflection about what values one wants to advance. By articulating the “values dimensions of our craft” (Greene, 2011, p. 8) and illustrating how these values are integrated into actual practice, meta-evaluative judgements about the quality of the evaluation can be made more readily. More importantly, other evaluators can learn about the challenges of employing different approaches within the contingencies of the real-world practice context.

The Thoughtful Evaluation model I developed and presented in Chapter Two resulted from a careful analysis of historical evaluation debates. It represents the values dimensions that I believe are critical to quality evaluation practice. I used this model as a vision to guide my practice, illustrated the attempts I made to employ it, and reflected on its suitability for the Project K evaluation context. Together these sections of the thesis provide useful insights that evaluators (especially those new to the field) need to be aware of. I offer the Thoughtful Evaluation model as a useful guide for evaluators who also share my commitment to the Community Psychology values of social justice, respect for human diversity, citizen participation and empowerment, a strengths-based and socio-ecological focus, and empirically-based social action.

LIMITATIONS AND LESSONS FOR PROGRAMME DEVELOPMENT

The high degree of consensus about Project K’s theory of change across organisational stakeholders makes evident that communication about how Project K is intended to operate has filtered down to the community-based practitioners. The logic model depiction of this theory of change can now serve as a communication tool so that this knowledge continues to be easily transferred to relevant stakeholder groups. Nevertheless, the Project K theory-building process presented opportunities to enhance a few programme delivery aspects.

Specifically, I noted four potential opportunities for programme development. These included: 1) difficulty engaging participants in the Community Challenge; 2) problems participants may have with reintegration into
their day to day lives after the Wilderness Adventure and Community Challenge; 3) increasing the importance of cultural consideration aspects in programme practice; and 4) making clear where physical health outcomes fit in Project K’s theory. These opportunities were described to FYD staff and Project K practitioners in a research dissemination session at a recent FYD conference and the workshop attendees worked in groups to make suggestions about how to improve each of these areas (see Table 1 in Appendix F). Here I reflect on some of their recommendations for action while also considering how the findings from the RCT study in Chapter Five can further inform these issues.

Several solutions put forward by the workshop attendees to address each of these opportunities fit with a theme that greater integration was needed with external support networks including families, schools, and the community. For instance, greater integration with community members, especially during the Community Challenge, should increase the likelihood of developing genuine relationships between young people and the adults in their communities. I indicated in Chapters Three and Four that this should increase the authenticity of service experiences and, consequently, participant engagement in this component. Furthermore, fostering cross-cultural relationships with community members and including community members who can share indigenous knowledge could address some of the cultural consideration issues. Greater family involvement could also contribute to this end as Asian, Pacific, and Māori culture all emphasise the importance of interconnectedness with family and collective notions of well-being (Ministry of Social Development, 2005; Sobrun-Maharaj & Wong, 2010; Wayne Francis Charitable Trust, 2011).

It was encouraging to see from the RCT findings that over the duration of the programme, differential effects did not occur across ethnic groups and that one year later the effects seemed to have had an even greater impact for young Asian participants (for social self-efficacy and academic achievement outcomes). However, Pacific youth were less likely to graduate from Project K and both Māori and Pacific participants were less likely to complete the evaluation surveys at post and one year post-programme. This raises concerns that their results may not be
adequately represented in this research or their needs may not have been adequately addressed during the programme and is linked to earlier points about the need to directly investigate any hidden perspectives associated with the potential iatrogenic effects. It was also clear that the Māori and Pacific youth still had needs with regards to academic success (relative to other ethnic groups) that were not being addressed by the programme. These issues support the greater need for cultural consideration within the programme. In addition, while Project K aims to enhance the holistic growth of a young person and thus is not an educational intervention per se, greater family and school involvement with the programme (especially when reintegrating back into life following the first two programme stages) could be beneficial in strengthening the programme’s impact on academic achievement outcomes.

In making the above recommendations, I strongly caution against any action towards greater integration with external networks without first reflecting carefully on the potential negative implications. For example, the mentoring literature indicates that there is a fine line between adequate and too much family involvement in the programme. Mentors are meant to provide individualised support and extensive family involvement can detract from the benefits of one-to-one attention (Morrow & Styles, 1995) – the focus should be on one young person, unless group mentoring aspects are incorporated. We also learned through the theory of change development process that some of the participants enter the programme having experienced disruptive life events, including dysfunction in their home lives and inappropriate family support. Thus additional complications could arise from involving other family members which would impact on the programme dynamic and require additional resources and support that may go beyond the current level of staff expertise.

As a final example to illustrate the need to consider integration with external networks carefully, I reiterate a lesson learned from Hollis et al.’s (2011) research. Young Māori participants spoke positively about the ethnic equality they experienced as part of their Project K journey. This gave them a rare opportunity to be individuals away from the negative stereotype associated with being Māori.
Programme designs would need to consider how to incorporate tikanga Māori principles into the programmes while still providing an atmosphere where young Māori can feel they are being treated like everyone else. Perhaps a focus on multiculturalism which includes the importance of the Te Tiriti o Waitangi/Treaty of Waitangi principles for New Zealand society is a good solution. When asked to provide feedback on cultural consideration, several programme development workshop attendees commented on the importance of multiculturalism within the programme, rather than a sole emphasis on biculturalism. As mentioned above, many also indicated a clear need for workforce training on cultural competence so that they could better attend to the participants’ unique cultural needs.

Greater integration of specific content across programme stages was another suggestion made by programme practitioners as a way to strengthen the impact on health and fitness. To clarify, the iterative but staged progression of learning appears to work very well for enhancing self-efficacy but Zhang (2011) found activities specifically addressing the physical health and lifestyle behaviours were quite disconnected (i.e. by eating healthy food in the wilderness, providing health-focused workshops in the Community Challenge, and working towards one health and fitness goal in the mentoring component). Likewise, the emphasis on academic outcomes seems to be isolated to choosing one academically-oriented goal during the mentoring component. If these two outcome categories are to remain primary Project K objectives then the programme would benefit from the inclusion of activities that repeatedly reinforce these themes across the programme stages, as is done for efficacy and other primary outcome categories, like connectedness.

At the same time, health and fitness and educational achievement outcomes were not commonly reported by the staff members I interviewed, by the Project K participants, or within the feedback on the logic model components from stakeholders who were not involved in the focus groups. This absence in the qualitative data makes me question whether Project K needs to emphasise these more in programme practice or less in programme documentation. It is important to recognise that Project K cannot address all well-being outcomes on its own. Additionally, the programme objectives would align more closely with the
programme’s theory of change if FYD clarified that improving health and fitness and academic achievement is secondary to increasing efficacy, connectedness, skill development, or motivation.

Even as secondary outcomes, the results for health and fitness (see Zhang, 2011) and NCEA achievement are a little disheartening; however, it must be said other programmes given “model” status have also had less than impressive results for some outcomes. For instance the Across Ages programme mentioned in youth development literature review (Chapter Three) has been given “model program” status (DuBois et al., 2001, p. 59) and is similar to Project K in that it is a multicomponent youth programme that combines a PYD curriculum with community service and mentoring during the school year. The effects from an Across Ages RCT evaluation were positive overall; however, the programme only had marginal effects relative to a Control group on general well-being and frequency of substance use. Furthermore, Aseltine, Dupre and Lamelin (as cited in DuBois et al., 2011) discovered than none of the positive effects from the Across Ages programme were sustained less than a year after programme completion. This demonstrates that positive effects across all outcome categories are difficult to obtain and sustain, even for “model” programmes.

Besides all of this, it is very encouraging to see that new developments initiated within FYD programmes even before dissemination of these findings fit well with some of the above recommendations. A new Community Development Strategy is being trialled in several different communities to streamline the primary school-based *Kiwi Can* programme, the Year 9 peer support Stars programme, and Project K within the same school communities, which will reinforce PYD lessons throughout middle childhood and adolescence. FYD is also working with other agencies within these communities to integrate support services. Moreover, Project K has been modified for delivery to *kura* (Māori immersion) school students. Cultural advisors have been involved in the process, programme designers have been refining programme strategies based on these pilot experiences, and similar initiatives have begun for programmes exclusive to Pacific students.
CLOSING REMARKS

In this day and age, evaluative activities can serve several different functions including formative assessments for programme improvement and knowledge-generation for wider spheres of social science research and practice (Chen, 1990; Donaldson, 2003, 2007; Patton, 1997; Rossi & Freeman, 1989; Weiss, 1997a, 1997b, 1998a, 1998b) but the central function of programme evaluation has always been to render a judgement about a programme’s merit or worth (Hall & Hall, 2004; House, 1993; Shadish, 1994; Shadish et al., 1991; Stake, 2004; D. Stufflebeam, 2001; Weiss, 1998a). This thesis was an evaluation of Project K and as such the end goal was to make a value judgement about the programme based on the information compiled. Overall, the majority of the evidence indicates that Project K adds considerable value to the lives of its participants, despite not appearing to deliver on all the outcomes it hopes to influence. I cannot comment on the cost-benefit aspects of Project K or its comparative value with other New Zealand-based youth development programmes but I do suspect the advanced and innovative way in which the best practice principles are integrated stands it above most others in the youth development field – overseas and in New Zealand.

At the same time, the comprehensive nature of this thesis provided me with opportunities to also contribute to the programme improvement and knowledge-generation functions. This was in large part due to my focus on the mechanisms and conditioning factors of programme success. In attempting to get “inside the white box” I was able to clarify several programme development opportunities which have important implications for researchers, policymakers, and practitioners beyond the local programme scene. The fifth principle of the Youth Development Strategy Aotearoa stipulates that “youth development needs good information” (Ministry of Youth Affairs, 2002, p. 8), yet, to date, there has been a paucity of such information. Considering the comprehensiveness and credibility (to both internal and external programme stakeholders), I offer that this research represents the kind of “good information” that is needed. Likewise, the delineation of the evaluation process from conceptual model, to actual practice details, to retrospective reflection is rare to see in evaluation research. In doing this I hope others can learn from my
mistakes and perhaps even be inspired to incorporate some of the ideas in their own evaluative endeavours.

Clearly much work remains to be done to find out the best ways to support young people and to close the evaluation research-practice gap. The fields of youth development and programme evaluation are growing at a fast pace in New Zealand. With this come many opportunities to produce research that can have a positive and substantial impact on social change. I very much look forward to being part of this future.
Appendix A: Ethical Documentation for Thesis
Memorandum of Understanding

Between the Foundation for Youth Development and the University of Auckland

This understanding covers the research to be conducted by Kelsey Deane as part of Study 1 of her PhD thesis through the Department of Psychology under the supervision of Niki Harré. Study 1 consists of an analysis of a secondary data set consisting of all evaluation measures collected to date for Project K (a youth development program supported by the Foundation for Youth Development) to identify patterns and relationships between and within variables.

The Foundation for Youth Development (FYD) and the University of Auckland agree that:

1. The research will be conducted as outlined in the ethics application titled *An investigation of factors related to improvement in self-efficacy* (Ref.# 2007/336) previously given to FYD. This application was approved by The University of Auckland Human Participants Ethics Committee on the 13th of September, 2007.

2. FYD agrees to allow Kelsey Deane access to all evaluation measures (both quantitative and qualitative) collected to date, however, identifying details of the participants will not be disclosed to Kelsey Deane for the purposes of Study 1.

3. Kelsey Deane will not copy any FYD materials or use them for any other purpose than to design, conduct and write up the research.

4. Kelsey Deane has full rights to use the data and results of the analysis to write her PhD thesis. The role of FYD in the design of the evaluation and gathering the data will be fully acknowledged.

5. Kelsey Deane and Niki Harré have a right to publish any findings from the data analysis in another outlet or talk about the results at a conference, with the media or in any other public forum after obtaining consensus from FYD regarding the way the research is presented. An appropriate person from FYD will be invited to take an authorship role, if this is allowed by the outlet concerned. In order to be an author on any studies, the person from FYD will need to be involved in the writing process. However, this will need to be done in a timely fashion so as not to impede the dissemination of the report. Where FYD is not involved as an author the presentation will acknowledge the source and contribution of FYD.
6. FYD have the right to use any findings resulting from the research for marketing and fundraising purposes upon reaching consensus with Kelsey Deane and Niki Harré regarding the way the research is to be presented. Any roles either Kelsey or Niki has had in the research process will be fully acknowledged in these communications.

7. Regardless of input from either party, oral and written reports of findings will be open, direct and honest in their disclosure, including limitations, in order to maintain the integrity of the research.

8. The report of findings from this analysis will be jointly owned by Kelsey Deane, Niki Harré and FYD/Project K.

9. Any official communication between the FYD, Project K and Kelsey Deane will be conducted via the FYD Research and Evaluation Coordinator (Julie Moore) and/or Executive Trustee (Jo-anne Wilkinson).

10. Regardless of any personnel changes within the FYD or those involved from the University of Auckland during the time of this study, the research will continue according to the guidelines outlined in this document.

Signed by:

__________________________ (PhD student/Primary Investigator)

__________________________ (Supervisor)

__________________________ (Head of Department of Psychology, Representing the University of Auckland)

__________________________ (Representative of FYD)

ORGANISATION INFORMATION SHEET

Experiences from the ground: Multiple stakeholders’ views about Project K and a self-reflective inquiry of the Thoughtful Evaluation Model.

To: The Foundation for Youth Development

Researchers: Kelsey Deane and Niki Harré

The Study

We have previously received your approval to work with the Project K evaluation data collected to date. We are now asking for the opportunity to approach any adult programme stakeholders (i.e. FYD and licensee staff members, trustees, teachers, principals, volunteers, Project K graduates over the age of 16, and/or their caregivers) with an invitation to become involved in the evaluation process and/or to obtain their perspectives concerning Project K. Research has demonstrated that an evaluation process that involves various stakeholders results in greater use of the findings. Furthermore, focusing only on the results of the standardised quantitative measures and the limited amount of qualitative information available in the current evaluation database will restrict our understanding of what is working well with Project K and why. Including various stakeholders’ experiences and perspectives as qualitative data will help to provide a more accurate and holistic understanding of Project K.

One major challenge in conducting programme evaluation research is obtaining research findings that are both credible to an academic community while still being useful and relevant to those directly involved with the programme. An ideal outcome of such research would be to achieve both. As was presented at a previous meeting with FYD staff, I (Kelsey) have recently developed a model to guide evaluation research which emphasizes a balance between features of credibility to an academic community (accuracy, systematic methods, objectivity) and an empowering relationship between researchers and programme stakeholders (inclusiveness, shared decisions, mutual respect) in hopes that this will increase the evaluation’s use by both groups.

Additionally, we are requesting your approval to examine our process as external evaluators in relation to the dimensions of this new model while working with Project K stakeholders over the next three years. We intend using the model I (Kelsey) have developed to guide our approach to the evaluation of Project K and would like to document and analyse whether our approach as evaluators is in accordance with the model previously presented. Many theories developed to guide evaluation research have yet to be tested empirically (by collecting data)
and much is unknown about which approaches work best in specific contexts. With your help we are hoping to get feedback regarding this new model. We hope our insights will be useful both to Project K and others doing evaluation research with youth programmes.

The Method

Qualitative Perspectives about Project K

Information about the programme discussed during meetings, forums, and informal conversations (for example, in person, via email, telephone, or postal communications) may also be documented and used as qualitative data to assess the programme(s) if the individuals involved in these interactions have consented to taking part in the study.

Self-reflective Inquiry into my Process as an Evaluator

I (Kelsey) will be using self-reflective inquiry to gather data about my process as an evaluator. This process is similar to participant observation and will involve documenting the proceedings of any interactions (related to my research with Project K) between me (Kelsey) and Project K stakeholders. This may also include interactions during meetings, telephone calls, emails or other discussions. Following these interactions I will reflect on my thoughts and feelings regarding the behaviour and comments produced by others in response to my evaluative approach, as well as my own behaviour and comments. I will be taking detailed notes of these reflections and these notes will be analysed for themes/narratives relating to the approach described in my evaluation model.

Upon completion of the first draft report of the findings obtained through these interactions, I will distribute the draft of the results to everyone who has participated in this study and will invite the opportunity for feedback on the observations made. This feedback may be incorporated into the final report as additional data.

Project K stakeholders may also be invited to take part in completing a short questionnaire following these interactions. This will involve answering questions which will require participants to reflect on their experiences, thoughts and feelings about the communication processes used.

Additionally, I am interested in hearing about Project K stakeholders’ previous experiences with research and/or evaluation (they will be asked not to disclose the name of any organisation or people that were involved). Learning about what worked or did not work previously for participants will help me develop an approach that is more suitable to the adult stakeholders involved in the programme. To gather this information, participants may be invited to take part in a short (15-30 minute) interview (either in person or by telephone). These interviews may be recorded by audiotape (if the participants have consented to this) to increase the quality of the interviewing process and to ensure detailed descriptions are obtained.

The Selection Process
The different Project K stakeholder groups are numerous and unfortunately it is not feasible to include everyone in the evaluation process. Therefore, the first step in the participant selection process is to determine which stakeholder groups have the most intimate knowledge about the programme and which groups are most likely to use the evaluation results. This will be done in collaboration with members of the Foundation for Youth Development National Support Office and Licensee Staff members.

Once the primary stakeholder groups have been identified an invitation to participate in the evaluation process will be sent to stakeholders within the relevant groups. Although we will do our best to accommodate all those that are interested in participating, we may need to select individuals within these stakeholder groups. We would, however, like to include members that are representative of several different levels within the organisation. Should new opportunities exist to explain the research and invite participation from new stakeholders, those individuals will be briefed and consent will be sought at that time.

Organisational Consent

In signing the Organisation Consent Form the Foundation for Youth Development grants permission to Kelsey Deane and Dr. Niki Harré to conduct this research and to approach any adult Project K stakeholders (again, this may include staff members, volunteers, trustees, teachers, principals, Project K graduates over the age of 16 and/or their caregivers). Although permission to participate at the organisational level is greatly appreciated, it is entirely voluntary and refusal to participate as an organisation will in no way affect other related research studies involving Kelsey Deane and/or Dr. Niki Harré.

Individual Participation

Obtaining signed consent from the Foundation for Youth Development to conduct this research will in no way oblige any individual Project K stakeholders to participate. Should an individual choose not to be involved in the self-reflective observational research, anything he or she says or does during these interactions will not be noted or analysed. Furthermore, if someone is concerned with anything he or she may have said or done (during observations, interviews, and/or in the questionnaire) or if he or she changes his or her mind about participating after having consented, he or she may ask the researcher to have that specific instance or all findings related to him or her removed from the data, as long as this is done within 30 days following the completion of the questionnaire, interview or observed interaction. This is to ensure timely dissemination of the final report. Participants will be reminded by email of their participation in this process on a six-monthly basis and their consent to continue participating in the study will be sought each year. Participants can withdraw from participation in this research at any time without having to give a reason and without penalty.

Anonymity

Only the researchers named on this sheet will have access to the observational field notes, interview transcripts and questionnaires. These will remain on our computers indefinitely; any hard copies will be destroyed within the next six years. Signed consent forms with the
participants' names will be kept locked in a filing cabinet separate from any associated data in Niki Harré's office at the university. Their responses will remain confidential and a pseudonym or participant code will be used in any research report. We will do our best to protect their identities but, please be aware that because these meetings may occur in small groups and because the administrative structure of Project K may be known to others, there is a chance that others may be able to guess which observations/responses are associated with some individuals. I do want to emphasise, however, that the purpose of this study is to investigate observations/experiences that reflect on our process as evaluators and on information that will help to refine Project K, not on other individuals' communication processes and/or performance.

Right to Veto

Due to the fact that Project K is a well-known programme, if Project K’s identity is concealed in any research report (which will be done if requested by the Foundation for Youth Development) the organisation may remain identifiable. To mitigate this, the Foundation for Youth Development will retain the right to veto any observations, results or conclusions made that are perceived to reveal Project K’s or related individuals’ identities for 30 days after receiving a draft of the results.

Dissemination of the Results

We will be providing a full report of the research findings related to this study to the Foundation for Youth Development and to any individuals who have participated in the study and have requested a copy of the report. We will also be providing the Foundation for Youth Development with a copy of the PhD thesis, which will include the current study. Decisions regarding the dissemination of the results to wider audiences will be made collaboratively with the Foundation for Youth Development National Support Office.

Thank you for your consideration in permitting adult Project K stakeholders to participate in this study. If you have any questions at all about your organisation’s participation in this study, or feel uncomfortable about any part of the research, or if you just have general questions about the research at any time I encourage you to contact:

Kelsey Deane on 3737599 Ext. 84204 or k.deane@auckland.ac.nz OR Dr Niki Harré on 3737599 Ext. 88512 or n.harre@auckland.ac.nz

The Head of the Psychology Department is Associate Professor Fred Seymour he can also be contacted on 373 7599 ext 88414 or f.seymour@auckland.ac.nz

For Ethical concerns contact: The Chair, University of Auckland Human Participants Ethics Committee, The University of Auckland, Room 005 Alfred Nathan House, 24 Princes Street, Private Bag 92019, Auckland. Tel: 373 7599 extn. 87830.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE for a period of four years from 18/02/2008 to 09/04/2012. Reference 2008/009.
ORGANISATIONAL CONSENT TO PARTICIPATE IN RESEARCH

This form will be held for a period of 6 years.

Principal Investigator: Kelsey Deane, Psychology Department
PhD Supervisor: Dr. Niki Harré, Psychology Department

Research Title: Experiences from the ground: Multiple stakeholders’ views about Project K and a self-reflective inquiry of the Thoughtful Evaluation Model.

- I have read the information provided and understand the explanation of this research.
- I have had the opportunity to ask any questions and have had my questions answered to my satisfaction.
- I have been assured that the Foundation for Youth Development or any individual participants may withdraw from this research project at any time without having to give any reasons and without penalty.
- I understand that the Foundation for Youth Development has the right to veto any data, observations, results or interpretations that may reveal Project K’s identity prior to inclusion in any final research report for 30 days after receiving the draft.
- I have been given a guarantee that the participants’ names will not be disclosed and that their responses will be anonymous, but I am aware a possibility exists that others could identify them.
- I agree that the researchers have the Foundation for Youth Development’s approval to conduct the research outlined in the Organisation Information Sheet and to approach any adult Project K stakeholders (this may include staff members, volunteers, trustees, teachers, principals, Project K graduates over the age of 16 and/or their caregivers) for participation.
- I understand that I will be given a copy of the final report.
- I understand that this consent form will be held until the 01 February 2014 when it will be destroyed.

Consent of the Organisation:

I ________________________ (Executive Trustee for the Foundation for Youth Development) give organisational consent to the research conducted by Kelsey Deane and Dr. Niki Harré examining the use of empowering processes in a collaborative relationship between researchers and other programme stakeholders.

Signed: ________________________ (Executive Trustee for the Foundation for Youth Development)

Date: ________________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE for a period of four years from 18/02/2008 to 09/04/2012. Reference 2008/009.
INDIVIDUAL PARTICIPANT INFORMATION SHEET

Experiences from the ground: Multiple stakeholders' views about Project K and a self-reflective inquiry of the Thoughtful Evaluation Model.

To: Project K Programme Stakeholders (adults)

Researchers: Kelsey Deane and Niki Harré

The Study

I (Kelsey Deane) am a PhD student from the University of Auckland. My supervisor, Dr. Niki Harré is a Senior Lecturer at the university. We will be conducting research related to Project K for approximately the next three years. One of the aims of the research is to analyse in detail the evaluation data gathered by licensee staff and the National Support Office over the past three years (thanks to some of you!) to get a better understanding of what is happening within the programme.

We feel that you have important information about the programme that could help us in understanding the answers to our evaluation questions in addition to the information the evaluation data can tell us. We would, therefore, like to hear your stories and opinions about the programme. We would also like to involve you in shaping the future evaluation of Project K.

One major challenge in conducting programme evaluation research is obtaining research findings that are both credible to an academic community while still being useful and relevant to those directly involved with the programme. An ideal outcome of such research would be to achieve both. I (Kelsey) have recently developed a model to guide evaluation research which emphasizes a balance between credibility (accuracy, systematic methods, objectivity) and an empowering relationship between researchers and programme stakeholders (inclusiveness, shared decisions, mutual respect) in hopes that this will increase the evaluation’s use by both groups.

This is the model we will use to guide our approach to the evaluation of Project K. Many theories created to guide evaluation research have yet to be tested empirically (by collecting data) and much is unknown about which approaches work best in specific contexts. With your help I am also hoping to get feedback about my approach as an evaluator in regards to this new model. Through this research we aim to develop understanding about our process as
external evaluators. We hope our insights will be useful both to Project K and others doing evaluation research with youth programmes.

The Method

I will be documenting any information about the programme discussed during meetings, forums, and informal conversations (for example, in person, via email, telephone, or postal communications) and may use this information as qualitative data to assess the programme if you have provided this information and have consented to taking part in the study.

I will also be using self-reflective inquiry to gather data about my process as an evaluator. This will involve documenting the proceedings of any interactions (related to my research with Project K) between you, other Project K stakeholders and me. This may also include interactions during meetings, telephone calls, emails or other discussions. Following these interactions I will reflect on my thoughts and feelings regarding the behaviour and comments produced by others in response to my evaluative methods, as well as my own behaviour and comments. I will be taking detailed notes of these proceedings and reflections and these notes will be analysed for themes/narratives relating to the approach described in my evaluation model.

Upon completion of the first draft report of the findings obtained through these interactions, I will distribute the draft of results to everyone who has participated in this study and will invite the opportunity for your feedback on the observations made. This feedback may be incorporated into the final report as additional data.

You may also be invited to take part in completing a short questionnaire following these interactions. This will take about 5 to 10 minutes and will involve answering questions which will require you to reflect on your experiences, thoughts and feelings about the collaborative evaluation processes used during that interaction.

Additionally, I am interested in hearing about Project K stakeholders’ previous experiences with research and/or evaluation (you will be asked not to disclose the name of any organisation or people that were involved). Learning about what worked or did not work previously for you will help me develop an approach that is more suitable to the adult stakeholders involved in the programme. To gather this information, you may be invited to take part in a short (15-30 minute) interview (either in person or by telephone). These interviews may be recorded by audiotape (if you have consented to this) to increase the quality of the interviewing process and to ensure detailed descriptions are obtained.

Voluntary Participation

Although your participation is greatly appreciated, it is entirely voluntary. We have previously obtained signed consent from the Foundation for Youth Development to conduct this research; however, this is in no way obliges you to participate. Should you choose not to be involved in the self-reflective observational research, anything you say or do during these interactions will not be noted or analysed. Furthermore, if you are concerned with anything you may have said or done (during observations, interviews, and/or in the questionnaire) or if you change your
mind about participating after having consented, you may ask the researcher to have that specific instance or all findings related to you removed from the data, as long as this is done within 30 days following the completion of the questionnaire, interview or observed interaction. This is to ensure timely distribution of the final report. You will be reminded by email of your participation in this process on a six-monthly basis and your consent to continue participating in the study will be sought each year. You can withdraw from participation in this research at any time without having to give a reason and without penalty.

You will not be able to access my (Kelsey’s) observational notes because this may include information about someone else, but you are welcome to ask me about the notes pertaining to you at anytime. You will be permitted access to your own interview transcript and/or recording to allow an opportunity to clarify the accuracy of your statements, and you will be given a copy of the audiotape with your interview if you have requested this on the consent form.

Anonymity

Only the researchers named on this sheet will have access to the observational field notes and questionnaires. These will remain on our computers indefinitely; any hard copies will be destroyed within the next six years. Signed consent forms with your names will be kept locked in a filing cabinet separate from any associated data in Niki Harré’s office at the university. Your responses will remain confidential and a pseudonym or participant code will be used in any research report. We will do our best to protect your identity but, please be aware that because these meetings may occur in small groups and because the administrative structure of Project K may be known to others, there is a chance that others may be able to guess which observations/responses are associated with you. I do want to emphasise, however, that the purpose of this study is to investigate observations/experiences that reflect on our process as evaluators and on the merit and worth of the programme, not on your individual performance or communication processes.

You are welcome to receive a final copy of the research findings regarding this study (please indicate this on the consent form), which will also be included in my PhD thesis along with other studies related to Project K. A full report will also be accessible from the Foundation for Youth Development National Support Office.

Thank you for your consideration in participating in this study. If you have any questions at all about your participation in this study, or feel uncomfortable about any part of the research, or if you just have general questions about the research at any time I encourage you to contact:

Kelsey Deane on 3737599 Ext. 84204 or k.deane@auckland.ac.nz OR Dr Niki Harré on 3737599 Ext. 88512 or n.harre@auckland.ac.nz

The Head of the Psychology Department is Associate Professor Fred Seymour he can also be contacted on 373 7599 ext 88414 or f.seymour@auckland.ac.nz
For Ethical concerns contact: The Chair, University of Auckland Human Participants Ethics Committee, The University of Auckland, Room 005 Alfred Nathan House, 24 Princes Street, Private Bag 92019, Auckland. Tel: 373 7599 extn. 87830.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE for a period of four years from 18/02/2008 to 09/04/2012. Reference 2008/009.
INDIVIDUAL CONSENT TO PARTICIPATE

This form will be held for a period of 6 years.

Principal Investigator: Kelsey Deane, Psychology Department
PhD Supervisor: Dr. Niki Harré, Psychology Department

Research Title: Experiences from the ground: Multiple stakeholders’ views about Project K and a self-reflective inquiry of the Thoughtful Evaluation Model.

- I have read the information provided and understand the explanation of this research.
- I have had the opportunity to ask any questions and have had my questions answered to my satisfaction.
- I have been assured that I may withdraw from this research project at any time without having to give any reasons and without penalty.
- I have been given a guarantee that my name will not be disclosed and my responses will be anonymous, but I am aware a possibility exists that others could identify me.
- I agree that the researchers may use the information I provide during interpersonal interactions between myself and Kelsey Deane as data to assess Project K and their process as evaluators.
- I understand that I will be given a copy of the final report if I request one.
- I understand that this consent form will be held until the 01 February 2014 when it will be destroyed.

☐ Please tick if you consent to being contacted to take part in the interview

☐ Please tick if you consent to having this interview recorded by audiotape

☐ Please tick if you would like an audiotape copy of your interview

☐ Please tick if you want a summary of the results. Email: ____________________

Consent of the Participant:

I ______________________ (participant name) give my consent to the research conducted by Kelsey Deane and Dr. Niki Harré examining the use of empowering processes in a collaborative relationship between researchers and other programme stakeholders.

Signed: ______________________ Date: ______________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE for a period of four years from 18/02/2008 to 09/04/2012. Reference 2008/009.
PARTICIPANT INFORMATION SHEET

Evaluation Questions

You are invited to share your opinions regarding which evaluation questions about Project K are most important to you, as a programme stakeholder by completing this short survey. We (Kelsey Deane, a PhD student and my supervisor, Dr. Niki Harré of the University of Auckland) are trying to find out which questions about Project K are the most important to different stakeholders.

Completing the questionnaire should only take a few minutes, though some people may take more or less time to do it. The questionnaire is totally anonymous so no individual’s responses can or will be identified. However, please be aware that because the administrative structure of Project K may be known to others and some stakeholder groups may be composed of small numbers, there is a chance that others may be able to guess which responses are yours.

The results of the votes will be presented to the Foundation for Youth Development and Project K licensees to demonstrate if and/or how these votes differ between different stakeholder groups within the Project K/Foundation for Youth Development organisation. These results may also be incorporated into my (Kelsey’s) PhD thesis and associated publications, but your identity will not be revealed.

Your participation is greatly appreciated and will help us to maximise the use of the evaluation results and recommendations by focusing our research on questions that are important to the majority of voting stakeholders; however participation is completely voluntary. You may withdraw from this project at any time up until you complete and submit your survey.

Your votes will be stored in an electronic database (or if completed as a hardcopy in a locked cabinet) in a secure room in the Psychology Department, and only Kelsey Deane, Dr. Niki Harré and their research associates will have access to the data. The hardcopies will be destroyed after 6 years, but the electronic versions will be stored indefinitely for research purposes.

For any questions regarding this project, please contact: Kelsey Deane or Niki Harré (details below). The Head of the Psychology Department is Associate Professor Fred Seymour. He can also be contacted on 373 7599 ext 88414 or f.seymour@auckland.ac.nz
For Ethical concerns contact: The Chair, University of Auckland Human Participants Ethics Committee, The University of Auckland, Room 005 Alfred Nathan House, 24 Princes Street, Private Bag 92019, Auckland. Tel: 373 7599 extn. 87830.

Kind Regards,

Kelsey Deane, Psychology Department, University of Auckland, HSB Building, Room 631. Phone: 09 373 7599, extn 84204. E-mail: k.deane@auckland.ac.nz

Dr. Niki Harré Psychology Department, University of Auckland, HSB Building, Room 537. Phone: 09 373 7599, extn 88512. E-mail: n.ha.re@auckland.ac.nz

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE for a period of four years from 18/02/2008 to 09/04/2012. Reference 2008/009.
PARTICIPANT INFORMATION SHEET

Evaluation Feedback

You are invited to share your opinions/feedback regarding some evaluation results about Project K. We (Kelsey Deane, a PhD student) and my supervisor, Dr. Niki Harré of the University of Auckland are trying to include multiple perspectives from various programme stakeholder groups in relation to the interpretation of the results obtained and/or issues of concern/ opinions about the evaluation process.

You will be asked to provide any feedback/opinions you may have regarding information about the Project K evaluation that will be presented in this workshop. Your feedback may be incorporated into related research reports; however, you will have the opportunity to indicate whether or not you would like your individual feedback to be incorporated as data in a report. You will also be asked to indicate where you fit within the structure of the Foundation for Youth Development organisation (your response to this question is optional).

The feedback is totally anonymous so no individual's responses can or will be identified. However, please be aware that because the administrative structure of Project K may be known to others and some stakeholder groups may be composed of small numbers, there is a chance that others may be able to guess which responses are yours. The summary of results will be made accessible to you via email if you include your email address at the end of the feedback form.

Your input is greatly appreciated and will help us to ensure that the interpretation of the evaluation results is representative of multiple perspectives and that the evaluation process is conducted in a manner that tries to accommodate different needs. However submitting any feedback is completely voluntary. You may terminate your participation in this workshop at any time and you may withdraw your submitted feedback any time up to two weeks following this workshop session by emailing Kelsey Deane (contact details provided below).

(please turn over)
Your feedback will be stored in an electronic database (and if completed as a hardcopy, in a locked cabinet) in a secure room in the Psychology Department, and only Kelsey Deane, Dr. Niki Harré and their research associates will have access to the data. The hardcopies will be destroyed by the project completion data (9/04/2012), but the electronic versions will be stored indefinitely for research purposes.

For any questions regarding this project, please contact: Kelsey Deane, Niki Harré (details below) or the Head of the Psychology Department, Dr. Douglas Elliffe, The University of Auckland, Private Bag 92019, Auckland. Phone 373 7599, extn 88557 or d.elliffe@auckland.ac.nz

For Ethical concerns contact: The Chair, University of Auckland Human Participants Ethics Committee, The University of Auckland, Room 005 Alfred Nathan House, 24 Princes Street, Private Bag 92019, Auckland. Tel: 373 7599 extn. 87830.

Kelsey Deane,
Psychology Department,
University of Auckland,
HSB Building, Room 642.
Phone: 09 373 7599, extn 84204.
E-mail: k.deane@auckland.ac.nz

Dr. Niki Harré
Psychology Department,
University of Auckland,
HSB Building, Room 537.
Phone: 09 373 7599, extn 8512.
E-mail: n.harré@auckland.ac.nz

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE on 13/02/2008 for a period of 4 years from 9/04/2008. Reference no 2008/009.
EVALUATION FEEDBACK

You will be presented with information that relates to the evaluation research about Project K being conducted by Kelsey Deane (a PhD student from the University of Auckland). You are welcome to include any comments, opinions and/or issues of concern regarding the information presented.

Please indicate whether or not you consent to the use of your feedback as data that may be presented in related research reports. All comments will remain anonymous and the identifying characteristics of any persons associated with the feedback will be removed.

○ Yes
○ No

THE PROJECT K PROGRAMME LOGIC MODEL
1. The Antecedent Condition

**Question:** What situation in your community is Project K designed to resolve?

*Feedback:*

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

2. The Student Profile

**Question:** Describe the typical PK student profile or common factors associated with the group of students selected for the programme?

*Feedback:*

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

3. The Project K Experiential Learning Cycle

**Question:** What are the most important processes occurring in the programme to produce desired outcomes?

*Feedback:*

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________
4. The Influencing Factors

**Question:** What are the key factors that would make this/these processes more likely to lead to positive results or would impede this from happening?

**Feedback:**

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________

5. The Outcomes

**Question:** What are the most important outcomes resulting from a student’s participation in the programme?

**Feedback:**

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________

Additional Comments?

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
To identify if there are any differences in responses based on the type or sector of employment within the Project K organisation please answer the following (optional):

I am a:

- Foundation for Youth Development National Support Office Staff Member
- Project K Licensee Staff Member
- Project K Programme Provider
- Trustee
- Volunteer
- Project K graduate
- Project K graduate caregiver
- Other __________________

To receive a summary of the results from this workshop please include your email below. If you choose to include your email, this page will be detached from the feedback provided above after 13/12/2012 to ensure any identifying characteristics of the person making the comments is not associated with the data.
Appendix B: Logic Model Materials
SEMI-STRUCTURED INTERVIEW SCHEDULE FOR PROJECT K
LOGIC MODEL DEVELOPMENT

1. Describe social situation is Project K designed to address? What prompted the decision to develop the programme?

2. Describe the typical type of students Project K is supposed to be reaching?

3. What are the most important processes within the Wilderness Adventure? [Try to avoid listing more than 3]

   a) Why is this process important?

   b) Take me through the sequence of effects that should typically follow from being involved in this process. What is typically the direct result of a student going through this process? How is this linked to Project K’s desired outcomes? How do you think this affects larger systems in the long run (family, school, community)? What factors may inhibit this process from occurring/ or what factors would make it more likely to happen?
4. What are the most important processes within the Community Challenge? [Try to avoid listing more than 3]

   a. Why is this process important?

   b) Take me through the sequence of effects that should typically follow from being involved in this process. What is typically the direct result of a student going through this process? How is this linked to Project K’s desired outcomes? How do you think this affects larger systems in the long run (family, school, community)? What factors may inhibit this process from occurring/ or what factors would make it more likely to happen?

5. What are the most important processes within the Mentoring component? [Try to avoid listing more than 3]

   a. Why is this process important?

   b) Take me through the sequence of effects that should typically follow from being involved in this process. What is typically the direct result of a student going through this process? How is this linked to Project K’s desired outcomes? How do you think this affects larger systems in the long run (family, school, community)? What factors
may inhibit this process from occurring/ or what factors would make it more likely to happen?

6. What are the most important outcomes arising from Project K?
   a. Immediate outcomes?
   b. Intermediate outcomes?
   c. Long-term outcomes?

7. Could you discuss what you think are the main barriers to programme expansion? What holds people back from running more programmes?

8. Could you discuss what you think is Project K’s most important unresolved issue?
Appendix C: Foundation for Youth Development Ethical Documentation for Project K Randomised Controlled Trial Evaluation
29th November 2012

University of Auckland

Dear Kelsey,

PhD Thesis Content Permission

The Foundation for Youth Development gives permission for Kelsey Deane to include the following documents in the appendices of her PhD thesis titled: “Project K in black & white: A theory-driven & randomised controlled trial evaluation of a youth development programme”, completed in partial fulfilment of the requirements for her PhD.

Form E8: Project K Information Sheet for School Principal
Form E9: Project K Consent Form for Principals
Form E10: Project K Information Sheet for Parents of Year 10 Students
Form E11: Project K Information Sheet for Year 10 students
Form E21: Project K Information Sheet for Parents of students selected for the evaluation
Form E23: Project K Consent form for parents of children selected for the evaluation
Form E22: Project K Information Sheet for students selected for the evaluation
Form E25: Project K General Consent Form for Students

Yours sincerely,

J Moore
Research & Evaluation Manager
Information Sheet for School Principal

You have already received detailed information about Project K and the students it targets. This is an information sheet to explain the procedures that will be used to select students into Project K and to conduct the evaluation of Project K programmes. It also explains how teachers, parents and students will be informed about Project K evaluation and selection.

To meet ethical guidelines for evaluation, we are required to provide comprehensive information sheets and consent forms for all participants as outlined below.

Selection into Project K – stage 1
To select students for Project K, we will be using a two-stage process. The first stage involves gathering information from the students and their teachers. As part of this, we will be asking all the students in Year 10 to fill in a self-efficacy questionnaire. The questionnaires will be delivered in class time during the week of (dates). They will be handed out by Project K staff. The questionnaire will have a code number on it, so no names will be used. We will have a list that matches names and code numbers. This list will only be retained until the selection procedures have been completed.

The questionnaire (attached) asks the students about how socially effective they feel, how effective they feel at doing school work, and how easily they can approach adults for help. It will take about 10-15 minutes to complete.

All teachers involved with Year 10 will receive an information sheet outlining the selection and evaluation procedures Project K will be using and explaining their role in detail (attached). We will also ask two teachers to fill in a set of questions for each student (attached) indicating the extent to which they feel the student is doing as well as he or she can socially, academically and in relating to adults. Ideally, we would like the two teachers who know the student best to do this. Teachers who agree to take part will be asked to sign a consent form (attached).

The students’ names will appear on the questionnaires we will ask teachers to complete, but the teachers’ names will not. Teachers will be identified by a code number that will be entered on the database to allow us to see if there are any differences in the responses between teachers. A list of teachers’ names and code numbers will be kept until stage 1 of the selection procedure is completed, when it will be destroyed. Each teacher’s participation is voluntary and teachers do not have to give a reason for not wishing to take part. Teachers may withdraw the information they have supplied within one week of the date we will give them for getting it to us. This date will be stated on the questionnaire.

The students’ self-efficacy responses and the teachers’ responses will be entered onto a database. Students are able to request to see their results. Any student who has viewed their results can request that they be corrected, if inaccurate.

© Project K 2005
We will use the information we receive from the students and the teachers to identify students that we feel could benefit most from Project K. To carry out the selection, we will use nine scores for each student. The first three scores will come from the student’s questionnaire and will reflect their academic, social and help-seeking self-efficacy. There will also be three scores from each teacher that target these dimensions. Students will be listed in order of those with the most scores out of the nine that are under the average for the school. Students will be selected from the top of the list, unless they are eliminated by the exclusion criteria. This process and the exclusion criteria will be explained to you verbally in more detail.

The students identified through these procedures will then enter the random selection process, described later in this information sheet. We will send a letter to all parents of these students. The letter will also give parents more information about Project K and the opportunity to decide if they want their child to enter the second selection process.

As we are only gathering this information for the purposes of selection into Project K, we will not be informing parents of their child’s results. However, if parents wish to find out the result they can contact us (details) by (date). After that date the list linking the code numbers on the database to students’ names will be destroyed. It will therefore not be possible to identify an individual student’s result from then. The information related to students who have been selected for the second stage of selection will be held in the Project K Evaluation Database. The selection data will be retained indefinitely, and the information may be used to help us verify our selection procedures.

Parents who do not wish their child to participate in this survey, or to have teachers fill in a set of questions about their child, will be asked to write to or phone either (Project K contact details) or (school contact details) by (date). If parents do not withdraw their child, we will invite the student to take part when we come to the school. Students who do not take part will be asked to do a quiet activity, such as read a book or do homework. Parents will also have the opportunity to withdraw the information provided by their child (date – one week after survey delivery), by contacting us or the school as described earlier in this paragraph.

Selection into Project K – Stage 2 and Evaluation Procedures

We are continuing to explore the effectiveness of Project K. To do this, we will offer the full Project K programme to some of the students identified by stage 1, and a control programme to other students. Some students may be reserves, in case we do not get the number of Project K or control students that we need. Parents of students identified in the first stage will be asked to sign a consent form if they would like their child to take part in Project K if they are offered the opportunity. They will also need to agree to be part of the evaluation, whether or not their child is offered a place in Project K, if they wish their child to take part in the second stage of the selection. Students themselves will be asked for verbal agreement to participate. Students who are eligible will then have their names put into a draw. We will draw names at random. This process means that some eligible students will be included in Project K and some will be included in the control programme.

© Project K 2005
We will continue to evaluate Project K and control students for a number of years, using the following measures. All measures are attached for your interest.

**Self-efficacy questionnaire.** This is filled in by all Year 10 students and asks about confidence in a number of areas; academic, social and relating to adults. **Self-reported health and lifestyle behaviours questionnaire.** This questionnaire asks Project K and control students to say how often they are involved in a number of activities. These include positive activities like sports and hobbies and risky activities like taking drugs and unsafe sexual practices. This questionnaire also asks about home and family environment. **Parent questionnaire.** This is a questionnaire we will ask parents to complete. It asks for impressions of the child’s behaviour. **Teacher questionnaire.** This is a questionnaire we will ask two teachers to complete about each student. It is very similar to the parent questionnaire. All teachers who agree to fill in these questionnaires will be asked to sign a consent form.

Parents, students and teachers will be asked to fill in these measures before Project K and the control programme begin, 18 months later when the programmes are completed, and again 1 year later. Students will also be asked to fill in the self-efficacy and health and lifestyle questionnaire in around 4-5 years from now. At that time, we will also ask each student questions about his or her life and experiences since school. The reason we want the measures filled in so often is that we want to see if Project K has an immediate effect on participants and if this effect is maintained as they leave school and enter training or the workforce. We estimate that it will take about 45 minutes of students’ time to fill in each set of questionnaires and it will take parents and teachers about 10 minutes to fill in the questionnaire about each child each time they are invited to do so.

We will also ask parent’s permission to obtain information on their child from the school. This includes PAT results or equivalent, NCEA results as they become available and notification that their child has left school. We will also ask them for permission for the school to give us a phone number and contact address for their child, so we can keep in contact with him or her. We would like to ask your permission for the school to give us this information regarding participants whose parents have agreed to this.

We may also want to contact the Ministry of Education to determine the student’s education status at the end of the programme, and one and three years post-programme.

The evaluation information we receive will be stored on a database that may include the participants’ names. This is for administrative purposes. The database will be set up so that only people directly involved in data entry and in management of the database will have access to completed questionnaires and associated names. This is limited to a small number of people who work through Project K National Support Office. The local Project K staff and teachers at the school will not be able to identify a particular student’s result. Participant’s names may also appear on the questionnaires, to ease delivery. However names will be removed after the data is entered and the questionnaires will be stored using a code number.
We will keep a list of the names of participants and their code numbers until the end of the evaluation, but this will be stored separately from the data itself and will only be available to Project K staff when necessary. It will not be possible to identify an individual in any reports of the results. No participants in the evaluation will be able to view questionnaires completed by someone else. For example, parents will not be able to view their child’s questionnaires, or the questionnaires the teachers complete on their child. Neither yourself nor teachers will not be able to view the parents or students questionnaires.

The data will be stored indefinitely at Project K National Support Office.

Parents will be informed that if, at some point in the future, they wish to stop participating in the evaluation, or withdraw their child, they are entitled to do so. They may also withdraw the data they or their child has supplied within two weeks of completing any particular questionnaire. Withdrawing themselves or their child from the evaluation once Project K has commenced will have no impact on their child’s participation in Project K or the control programme. They will also be informed it will have no impact on their child’s schooling.

Once a student turns 16, parents will no longer be able to withdraw them or their data, although the student will still be able to withdraw him or herself. At this stage, a student who has been withdrawn from the evaluation by a parent but is still participating in Project K or the control programme may be invited to rejoin the evaluation.

Parents will be informed that they do not have to give a reason for withdrawing, however it would be very useful to us if they did so.

Teachers will be identified by a code number. Only the code number will be entered on the database. We will retain a list linking the teachers’ names and code numbers for the duration of the school phase of the evaluation, which is up to four years. It is important for us to be able to distinguish different teachers, mostly for statistical purposes. However, the list linking code numbers and names will only be available to Project K staff, and teachers will never be identified personally at any stage of the analysis or in any reports written.

Questionnaires will usually be completed by the Project K and control programme participants in school time, in a group setting supervised by someone from Project K. Participants do not have to answer all questions, or give a reason for missing some questions, however we will encourage students to fully complete each questionnaire by asking them to check their answers and fill in any questions they left out by mistake.

Parents will be asked to complete the questionnaires on their child either at home or at a group meeting. Teachers will be asked to complete the questionnaires at school. Parents and teachers do not have to answer all the questions, but we will ask that they please try and do this unless they have a particular objection to some items. It is very useful for us if we have complete responses from everyone. You may withdraw the school from the evaluation at any stage prior to the collection of all data. This means we would cease to do any data collection at the
school, or invite the teachers to continue to be involved. However, if you did this, parents and students who are participating in the evaluation may still be asked to take part outside of the school. You will not be able to withdraw any of the data once it has been obtained from teachers, parents or students.

If you are at all concerned about the evaluation at anytime, or anyone participating in the evaluation expresses concerns to you, we strongly encourage you to contact us. Our contact details appear at the end.

This evaluation has been funded by Project K Trust, the Ministry of Social Development and The Southern Trust. The Ministry also contributes funding for the programmes which are the subject of this evaluation process.

The results of data collected from the national evaluation of Project K will be reported in a number of ways, and will probably be published in academic journals, local media and on the internet. We will send a report to you, and can send one to any other interested people. Note that it may be several years before we can provide a final report.

If you have any questions you may contact the following people:
Consent Form for Principals

I have received and read an information sheet about Project K selection and evaluation and had the opportunity to ask questions.

I agree that my school participate in the evaluation of Project K.

I understand that I may withdraw the school from the evaluation at any stage prior to the collection of all data. I understand that if I do this, parents and students who are participating in the evaluation may still take part outside of the school.

I understand that I may not withdraw any information once it has been obtained from students, teachers or parents.

I understand that the data will be kept indefinitely.

I agree that the school will identify someone that parents and students can contact if they wish to withdraw themselves or their child from the Project K selection process or evaluation.

I agree that the school will identify someone that parents, students and teachers can contact if they have questions about the evaluation. This person will pass on any questions he or she cannot answer to the Project K staff.

I agree that the school will identify a process by which parents can obtain the results of their child’s initial self-efficacy questionnaire and the teachers’ responses about their child obtained during stage 1 of the selection. I understand that this information will only be available to parents for the period that is stated on the parent information sheet.

I agree that Project K may obtain the following information from the school, on participants whose parents have consented to this: NCEA results, date of leaving school, contact phone number and address.

I agree that Project K may obtain the following information from the Ministry of Education: education status at end of programme, and one and three years post-programme.

Name:

Name of school:

Date:

Signed:

© Project K 2005
(Name of school) participates in a youth development programme called Project K. Project K consists of three parts; a Wilderness Adventure, a Community Challenge and a Mentoring Partnership. The programme aims to help students achieve their full potential. It is targeted at students who may be underperforming in some areas and could potentially benefit from the approach used by Project K. As a first step in selecting students into the programme, we would like to ask all Year 10 students at the school to fill in a self-efficacy questionnaire. The questionnaires will be delivered in the week of (dates). They will be handed out by Project K staff. The questionnaires will have code numbers, so no names will be used. We will have a list that matches students' names and code numbers. This list will only be kept until the selection procedures have been completed.

The questionnaire asks the students about how socially effective they feel, how effective they feel at doing school work, and how easily they can approach adults for help. It will take about 10-15 minutes to complete. We will also ask two teachers to fill in a set of questions for each student indicating the extent to which they feel the student is doing as well as he or she can socially, academically and in relating to adults. The students’ self-efficacy questionnaires and the teachers’ responses will be entered on a database. The database will only be available to Project K staff and the information on it will not be used for other purposes in the school.

We will use the information we receive from the students and the teachers, as well as meetings with selected school staff, to identify students that we feel could benefit most from Project K. (If you wish to know more about these procedures please contact us.) These students will then enter a second selection process. If your child is identified, we will be sending you a letter by (date). This letter will also give you more information about Project K. You will have the opportunity at that time to decide if you want your child to enter the second selection process, which will be explained in detail.

As we are only gathering this information for selection into Project K, we will not be informing parents of their child’s results. However, if you wish to find out your child’s results you may do so by contacting (details) by (date). After that date, the information on children who have been selected will be stored in the evaluation database. It will not be possible for anyone other than Project K evaluation staff to identify your child’s result from that date. The database will be retained indefinitely, and the information may be used to help us verify our selection procedures.

If you do not wish your child to participate in this survey, or to have teachers fill in a set of questions about your child, please write to or phone either (Project K contact details) or (school contact details) by (date). If we do not hear from you, we will invite your child to take part when we come to the school. Students who do not take part will be asked to do a quiet activity, such as read a book or do homework. If you wish to withdraw the information provided by your child you may...
do so until (date – one week after survey delivery), by contacting us or the school as described in this paragraph.
(Name of school) takes part in a programme for young people called Project K. Project K has three parts; a Wilderness Adventure, a Community Challenge and a Mentoring Partnership. There are a limited number of places and we can’t promise you will be selected. But as a first step to help choose students for the programme, we would like you to fill in a questionnaire. The questionnaire will have a code number on it, so no names will be used. We will have a list that matches names and code numbers. We will destroy that list after the Project K students have been chosen.

The questionnaire asks you how you feel about yourself. This includes how you feel about yourself socially, how you feel about your school work, and how easy it is for you to get help from adults. It will take you about 10-15 minutes to fill out. We will also ask two teachers some questions about how they think you are getting on socially, with your school work and with getting help from adults. Your questionnaire and your teachers’ answers about you will be put onto a database. This database will include your name, as it makes it easier for us to choose Project K students. However, only Project K staff will see the database. The school won’t use it for anything else.

We will use the information on the database, and other information from your school, to choose students we feel could get the most out of Project K. These students will then enter a second selection step. If you are one of those students, we will send your parents a letter by (date).

We won’t be telling you the result of your questionnaire or the teachers’ answers about you. Your parents can get the result from (details) as long as they do it by (date). You can request that your results be corrected if you believe they are inaccurate. After that date, the results and names of the students who have been chosen to go through to the second stage of the selection will be stored in the Project K evaluation database.

If you do not want to fill in the questionnaire, you don’t have to. You can either just not take a questionnaire when we hand it out, or not fill it in. You don’t have to answer all the questions, but it is very helpful if you do. If you miss some out, this may make it harder for us to decide if Project K might be helpful for you. If you do not take part you will be asked to do a quiet activity, such as read a book or do homework. If you decide you don’t want us to use your questionnaire after you’ve filled it in, you can get us to destroy it, and remove it from the database by phoning us (contact details) or the school (contact details) by (date – one week after survey delivery).

© Project K 2005
Your child has been identified as someone who may benefit from Project K.

We are continuing to explore the effectiveness of Project K through evaluation. As part of the evaluation, we have identified all the Year 10 students from your child’s school who may benefit. We will offer the full Project K programme to some of these students and a control programme to other students. Some students may be reserves, in case we do not get the number of Project K or control students that we need. All students who agree, and whose parents agree, that they would like to do Project K if offered the opportunity and who agree to be part of the evaluation, will have their names put into a draw. We will then draw names at random. This process means that some eligible students will be included in Project K and some will be included in the control programme.

We would like to invite you and your child to be part of the evaluation of Project K. This will involve including your child in the selection process, described above. Please note that we can only include your child in the selection process if you agree that your child will be part of the evaluation of Project K, whether they are selected for Project K or the control programme.

The evaluation will continue for a number of years and involves the following measures. You can get more information about these from the Programme Director, or from the people listed at the bottom of this information sheet. It is possible for you to view the measures by arranging this with the Programme Director.

**Self-efficacy questionnaire.** This is filled in by your child and asks about her or his confidence in a number of areas: academic, social and relating to adults.

**Self-reported health and lifestyle behaviours questionnaire.** This questionnaire asks your child to say how often he or she is involved in a number of activities. These include positive activities like sports and hobbies and risky activities like taking drugs and unsafe sexual practices. This questionnaire also asks about home and family environment.

**Parent questionnaire.** This is a questionnaire we will ask you to complete. It asks your impressions of your child’s behaviour.

**Teacher questionnaire.** This is a questionnaire we will ask two teachers to complete about your child. It is very similar to the questionnaire you will be asked to complete.

You, your child and your child’s teachers will be asked to fill in these measures before Project K and the control programme begin, 18 months later when the programmes are completed, and again 1 year later. Your child will also be asked to fill in the self-efficacy and health and lifestyle questionnaire in around 4-5 years from now. At this time, we will also ask your child other questions about his or her life and experiences since school. The reason that we want the measures filled in
so often is that we want to see if Project K has an immediate effect on participants and if this effect is maintained as they leave school and enter training or the workforce. We estimate that it will take about 45 minutes of your child’s time to fill in each set of questionnaires and it will take you about 10 minutes to fill in the questionnaire about your child each time you are invited to do so.

Questionnaires will usually be completed by the Project K and control programme participants in school time, in a group setting supervised by someone from Project K. Participants do not have to answer all questions, or give a reason for missing some questions, however we will encourage students to fully complete each questionnaire by asking them to check their answers and fill in any questions they left out by mistake. You will be asked to complete the questionnaires on your child either at home or at a group meeting. You do not have to answer all the questions, but we would ask that you please try and do this unless you have a particular objection to some items. It is very useful for us if we have complete responses from everyone.

We would also like to ask your permission to obtain information on your child from the school. We would like to ask the school for your child’s NCEA results as they become available and for them to tell us when your child leaves school. If necessary we would like the school to be able to give us a phone number and contact address for your child, so that we can keep in touch with him or her.

We may also want to contact the Ministry of Education to determine your child’s education status at the end of the programme, and one and three years post-programme.

The evaluation information we receive from your child will be stored on a database that may include your child’s name. This is for administrative purposes. The database will be set up so that only people directly involved in data entry and in management of the database will have access to your child’s name. This is limited to a small number of people who work through Project K National Support Office. The local Project K staff and teachers at the school will not be able to identify a particular child’s result. Your child’s name may also appear on the questionnaires, to ease delivery. However names will be removed after the data is entered and the questionnaires will be stored using a code number. We will keep a list of the names of participants and their code numbers until the end of the evaluation, but this will be stored separately from the data itself and will only be available to Project K staff when necessary. It will not be possible to identify you or your child in any reports of the results. No participants in the evaluation will be able to view questionnaires completed by someone else. This means for example, that you will not be able to view your child’s questionnaires, or the questionnaires the teachers complete on your child.

The data will be stored indefinitely at Project K National Support Office.

If you want your child to be entered into the random selection process for Project K described earlier, you need to agree to you and your child being part of the evaluation. It is very important for us that we can gather information on all the participants to judge how effective Project K is. However, if at some point in the
future you wish to stop participating in the evaluation, or withdraw your child, you
are entitled to do this. You may also withdraw the data you or your child has
supplied within two weeks of completing any particular questionnaire. Withdrawing
yourself or your child from the evaluation once Project K has commenced will have
no impact on your child’s participation in Project K or the control programme. It will
also have no impact on your child’s schooling.

Please note, that once your child turns 16, you will no longer be able to withdraw
your child, or your child’s data, although your child will still be able to withdraw him
or herself. At this stage, a child who has been withdrawn from the evaluation by a
parent but is still participating in Project K or the control programme may be invited
to rejoin the evaluation.

You do not have to give a reason for withdrawing, however it would be very useful
to us if you did so. If you are at all concerned about the evaluation at any time, we
strongly encourage you to contact us. Our contact details appear at the end.

This evaluation has been funded by Project K Trust, the Ministry of Social
Development and The Southern Trust. The Ministry also contributes funding for the
programmes which are the subject of this evaluation process.

The results of data collected from a national evaluation of Project K will be reported
in a number of ways, and will probably be published in academic journals, local
media and on the internet. We will be sending a report to the principal of your
school, and can send you a report too. Please indicate your interest in this on the
form provided. Note that it may be several years before we can provide a final
report.

If you have any questions you may contact the following people:
Consent form for parents of children selected for the evaluation

I have received an information sheet and had the opportunity to ask questions.

I agree that myself and my child will take part in the evaluation of Project K. I agree to this, regardless of whether my child is selected for Project K or the control programme.

I understand that I may withdraw the questionnaires supplied by myself or my child within a week of completing them.

I understand that withdrawing myself or my child from the evaluation will have no impact on my child’s participation in Project K or the control programme. It will also have no impact on my child’s schooling.

I understand that once my child is 16 years old, I may no longer withdraw him or her or his or her data, although my child will be able to withdraw him or herself. At this stage, a child who has been withdrawn from the evaluation by a parent but is still participating in Project K or the control programme may be invited to rejoin the evaluation.

I understand that the data will be kept indefinitely.

I agree that Project K may obtain the following information on my child from the school: NCEA results, date of leaving school, contact phone number and address.

I agree that Project K may obtain the following information from the Ministry of Education: education status at end of programme, and one and three years post-programme.

Name:

Child’s name:

Relationship to child:

Ph (Home):

Ph (Work):

Ph (Mobile):

Date:

Signed:

© Project K 2005
Information sheet for students selected for the
evaluation

You are one of the students from Year 10 that we think may benefit from Project K.

We are exploring how well Project K works, and we are trying to find this out by doing a study. For our study, some of the students we think could get help from Project K will take part in Project K, and some of these students will take part in another shorter programme, called a control programme. We will then compare the students who do Project K with the control programme students, looking at how they feel about themselves and how their teachers and parents think they are doing. This will help us find out if Project K is useful for young people.

If you and your parents agree that you would like to be part of Project K or the control programme, we will put your name in a hat. Names will be drawn out for the Project K programme and for the control programme. Some people’s names may not get drawn at all, and they will be reserves. You can’t choose which programme you get into. It is just a matter of luck.

It will take us a number of years to know how useful Project K is, and so if you get into Project K or the control programme, we would like to ask you, your parents and your teachers some questions about you at least once a year for four years. It will take you about 45 minutes each time to fill in our questionnaires. You will usually do this at school.

When you get a questionnaire, you do not have to answer all questions, or give a reason for missing some questions, however we will encourage you to check your questionnaire at the end and fill in any questions you have left out by mistake. It is very useful for us if you answer every question.

We will also ask your parent’s permission to let us ask the school for your NCEA results, and for the school to tell us when you leave. We will ask your parents if the school can give us a phone number and contact address for you if you move, so that we can keep in touch.

We may also want to contact the Ministry of Education to determine your education status at the end of the programme, and one and three years post-programme.

Your answers to the questionnaires will be kept on a computer database that may hold your name. This is so we can easily enter the information about you. The database will be set up so that only people from Project K National Support Office can see your name. No one from school will be able to see your name or know which answers are yours. Your parents will not be able to see your name or know which answers are yours. Even you will not be able to do this. Your name might be on the questionnaires you fill out, but it will be cut off after we’ve entered your answers on the computer. When we tell the school about the results of our study on Project K, or publish anything about the study, there is no way for anyone to
know who you are.

You can ask to see your results. Any student who has viewed their results can ask that they be corrected, if inaccurate.

If you decide that you want your name to go in the hat for Project K or the control programme, you can still change your mind later. If you get into Project K and then don’t want to be part of the study to find out how useful Project K is, you can say this. You don’t have to tell us why, and you will still be allowed to stay in Project K, even if you drop out of the study. However, it is very helpful if you stay in the study. This study is the best way for us to find out if Project K is useful.

Until you are sixteen, your parents can say at any time if they don’t want you to be in the study. Once you are sixteen they can’t do this any more. You can decide at any time, either now or after you are sixteen to stop taking part.

If you are worried about this study or have any questions, we strongly encourage you to contact us. Our contact details appear at the end.

This evaluation has been funded by Project K Trust, the Ministry of Social Development and The Southern Trust. The Ministry also contributes funding for the programmes which are the subject of this evaluation process.

The results of data collected from a national evaluation of Project K may be reported in a number of ways, for example, in university journals, in the paper, and on the internet. We will be sending a report to the principal of your school, and can send you a report too. Please say if you want a report on the form. Note that it may be several years before we can provide a final report.

If you have any questions you can contact the following people:
General Consent Form for Students

I have received an information sheet about Project K evaluation and had the opportunity to ask questions.

I agree to take part in the evaluation of Project K. I agree to this, regardless of whether I am selected for Project K or the control programme.

I understand that I may ask to see my results and have them corrected if I believe them to be inaccurate.

I understand that I may withdraw my questionnaires within a week of completing them.

I understand that withdrawing my data will have no impact on my participation in Project K or the control programme. It will also have no impact on my schooling.

I understand that once I am 16 years old, my parent/s may no longer withdraw my data, although I will be able to withdraw my own data. If I have been withdrawn from the evaluation by my parent/s, I may be invited to rejoin the evaluation at this time.

I understand that my data will be kept indefinitely.

I agree that Project K may obtain the following information about me from my school: NCEA results, date of leaving school, contact phone number and address.

I agree that Project K may obtain the following information from the Ministry of Education: education status at end of programme, and one and three years post programme.

Name:

School:

Ph (Home):

Ph (Mobile):

Date:

Signed:

© Project K 2005
Appendix D: Foundation for Youth Development Self-Efficacy Questionnaires
29th November 2012

University of Auckland

Dear Kelsey,

PhD Thesis Content Permission

The Foundation for Youth Development gives permission for Kelsey Deane to include the following documents in the appendices of her PhD thesis titled: "Project K in black & white: A theory-driven & randomised controlled trial evaluation of a youth development programme", completed in partial fulfilment of the requirements for her PhD.

Form E13: Project K Self-Efficacy Selection Survey
Form E438: Project K One Year Post-programme Survey (School students)
Form E438b: Project K One Year Post-programme Survey (School leavers)

Yours sincerely,

[Signature]

J Moore
Research & Evaluation Manager
DATE:______________

CODE:______________

SCHOOL:_____________________________________

YEAR IN SCHOOL:______________

CLASS:______________

Please tick one circle

MALE: O
FEMALE: O

ETHNICITY/CULTURE: (Please tick the circle for the ethnic group(s) you belong to. You may tick more than one)

O NZ Maori  O NZ European/Pakeha
O Tokelauan  O Other European
O Fijian  O South-East Asian
O Niuean  O Indian
O Tongan  O Chinese
O Cook Islands Maori  O Other Asian (e.g. Japanese, Korean)
O Samoan  O Other (e.g. African, South American)
O Other Pacific Islands

A code number is used to protect your privacy.
This is not a test. There are no right or wrong answers.
Your opinions are important to us. Your answers will help guide the development of a programme which is designed to help young people.

It is very helpful if you please answer every question. You will be asked to check your questionnaire to make sure it is complete before handing it in.

Please rate how well you can do the things below by ticking one circle for each question, using a pencil or black pen. ‘6’ means you can do it very well and ‘1’ means you cannot do it at all well. We are interested in your honest response. There are no right or wrong answers.

© Project K 2005
1. How well can you study when there are other interesting things to do?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

2. How well can you express your opinions when your classmates disagree with you?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

3. How well can you study for a test?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

4. How well can you become friends with other people?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

5. How well can you get adults to help you with a problem?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

6. How well can you succeed in finishing all your homework?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

7. How well can you have a chat with an unfamiliar person of your age?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
8. How well can you pay attention during class?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

9. How well can you cooperate with your classmates?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

10. How well can you get the information you need from adults?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

11. How well can you remember information given in class?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

12. How well can you take part in class discussions?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

13. How well can you motivate yourself to do school work?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

14. How well can you participate in class activities?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

15. How well can you get teachers to help you when you get stuck on school work?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
16. How well can you finish school assignments?

Not well at all          Quite well          Very well
1                   2                   3                   4                   5                   6
O                   O                   O                   O                   O                   O

17. How well can you succeed in staying friends with other people?

Not well at all          Quite well          Very well
1                   2                   3                   4                   5                   6
O                   O                   O                   O                   O                   O

18. How well can you succeed in satisfying your teachers with your school work?

Not well at all          Quite well          Very well
1                   2                   3                   4                   5                   6
O                   O                   O                   O                   O                   O

19. How well can you work in a group?

Not well at all          Quite well          Very well
1                   2                   3                   4                   5                   6
O                   O                   O                   O                   O                   O

20. How well can you get school staff to help you, when you have a problem at school?

Not well at all          Quite well          Very well
1                   2                   3                   4                   5                   6
O                   O                   O                   O                   O                   O

21. Do you have any goals for the year ahead?

No  O

Yes  O If you answered “yes”, please write the three most important ones below. They could be related to school, out of school activities, or personal goals. If you have less than three goals, write down the one or two you do have.

1. __________________________________________________________
2. __________________________________________________________
3. __________________________________________________________

PLEASE CHECK THAT YOU HAVE ANSWERED ALL THE QUESTIONS. THANK YOU

© Project K 2005
One Year Post-programme Survey
(School students)

DATE: ______________________

CODE: ______________________

SCHOOL: ______________________

YEAR IN SCHOOL: ____________

CLASS: ______________________

Please tick one circle

MALE: O
FEMALE: O

ETHNICITY/CULTURE: (Please tick the circle for the ethnic group(s) you belong to. You may tick more than one.)

- O NZ Maori
- O Tokelauan
- O Fijian
- O Niuean
- O Tongan
- O Cook Islands Maori
- O Samoan
- O Other Pacific Islands

- O NZ European/Pakeha
- O Other European
- O South-East Asian
- O Indian
- O Chinese
- O Other Asian (e.g. Japanese, Korean)
- O Other (e.g. African, South American)

A code number is used to protect your privacy.
This is not a test. There are no right or wrong answers.
Your opinions are important to us. Your answers will help guide the development of a programme which is designed to help young people.

It is very helpful if you please answer every question. You will be asked to check your questionnaire to make sure it is complete before handing it in.
Please rate how well you can do the things below by ticking one circle for each question, using a pencil or black pen. ‘6’ means you can do it very well and ‘1’ means you cannot do it at all well. We are interested in your honest response. There are no right or wrong answers.

© Project K 2005
1. How well can you study when there are other interesting things to do?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

2. How well can you express your opinions when your classmates disagree with you?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

3. How well can you express your opinions when people your own age disagree with you?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

4. How well can you study for a test?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

5. How well can you become friends with other people?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

6. How well can you get adults to help you with a problem?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

7. How well can you succeed in finishing all your homework?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
8. How well can you have a chat with an unfamiliar person of your age?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

9. How well can you pay attention during class?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

10. How well can you cooperate with your classmates?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

11. How well can you cooperate with people your own age?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

12. How well can you get the information you need from adults?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

13. How well can you remember information given in class?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

14. How well can you take part in class discussions?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

15. How well can you take part in group discussions?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
16. How well can you motivate yourself to do school work?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

17. How well can you participate in class activities?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

18. How well can you participate in group activities?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

19. How well can you get teachers to help you when you get stuck on school work?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

20. How well can you finish school assignments?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

21. How well can you succeed in staying friends with other people?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

22. How well can you succeed in satisfying your teachers with your school work?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

23. How well can you work in a group?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

© Project K 2005
24. How well can you get school staff to help you, when you have a problem at school?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>O</td>
</tr>
</tbody>
</table>

Thinking about your future career/job options, please read each statement carefully and indicate how much confidence you have that you could accomplish each of these tasks.

25. Work out what job would be best for you

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

26. Decide what you are looking for in a job

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

27. Find information about further education/training courses that will help you get the job you want

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

28. Prepare a good resume/CV

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

29. Perform well in a job interview

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

30. Choose a further education/training course or a job that will fit your interests

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
31. Determine the steps you need to take to successfully complete your chosen further education/training course

No confidence at all  Very little confidence  Moderate confidence  A lot of confidence  Complete confidence

32. Identify employers and further education/training institutions relevant to your career path

No confidence at all  Very little confidence  Moderate confidence  A lot of confidence  Complete confidence

Thank you. Please check that you have answered all the questions.
One Year Post-programme Survey
(School leavers)

Licensees to complete:

DATE:

CODE:

PROGRAMME NAME:

Please tick one circle

MALE:  O
FEMALE: O

ETHNICITY/CULTURE: (Please tick the circle for the ethnic group(s) you belong to. You may tick more than one.)

O NZ Maori          O NZ European/Pakeha
O Tokelauan         O Other European
O Fijian            O South-East Asian
O Niuean            O Indian
O Tongan            O Chinese
O Cook Islands Maori O Other Asian (e.g. Japanese, Korean)
O Samoan            O Other (e.g. African, South American)
O Other Pacific Islands

A code number is used to protect your privacy.
This is not a test. There are no right or wrong answers.
Your opinions are important to us. Your answers will help guide the development of a programme which is designed to help young people.

It is very helpful if you please answer every question. You will be asked to check your questionnaire to make sure it is complete before handing it in.

© Project K 2008
Please rate how well you can do the things below by ticking one circle for each question, using a pencil or black pen. ‘6’ means you can do it very well and ‘1’ means you cannot do it at all well. We are interested in your honest response. There are no right or wrong answers.

1. How well can you express your opinions when people your own age disagree with you?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

2. How well can you become friends with other people?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

3. How well can you get adults to help you with a problem?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

4. How well can you have a chat with an unfamiliar person of your age?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

5. How well can you cooperate with people your own age?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

6. How well can you get the information you need from adults?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

7. How well can you take part in group discussions?

   Not well at all | Quite well | Very well
   1 | 2 | 3 | 4 | 5 | 6
   O | O | O | O | O | O

© Project K 2008
8. How well can you participate in group activities?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

9. How well can you succeed in staying friends with other people?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

10. How well can you work in a group?

<table>
<thead>
<tr>
<th>Not well at all</th>
<th>Quite well</th>
<th>Very well</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

Thinking about your future career/job options, please read each statement carefully and indicate how much confidence you have that you could accomplish each of these tasks.

11. Work out what job would be best for you

<table>
<thead>
<tr>
<th>No confidence</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>at all</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

12. Decide what you are looking for in a job

<table>
<thead>
<tr>
<th>No confidence</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>at all</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

13. Find information about further education/training courses that will help you get the job you want

<table>
<thead>
<tr>
<th>No confidence</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>at all</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

14. Prepare a good resume/CV

<table>
<thead>
<tr>
<th>No confidence</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>at all</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
15. Perform well in a job interview

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

16. Choose a further education/training course or a job that will fit your interests

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

17. Determine the steps you need to take to successfully complete your chosen further education/training course

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

18. Identify employers and further education/training institutions relevant to your career path

<table>
<thead>
<tr>
<th>No confidence at all</th>
<th>Very little confidence</th>
<th>Moderate confidence</th>
<th>A lot of confidence</th>
<th>Complete confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

Thank you. Please check that you have answered all the questions.
### Appendix E: Supplementary Analyses for Chapter Five
### Randomised Controlled Trial Study

Table 1. Principal Components Analysis of Pre-Programme Self-Efficacy Items with Varimax Rotation (Eigenvalues over 1)

<table>
<thead>
<tr>
<th>Items</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Academic self-efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>How well can you succeed in finishing all your homework? A3</td>
<td>.78</td>
</tr>
<tr>
<td>How well can you finish school assignments? A7</td>
<td>.77</td>
</tr>
<tr>
<td>How well can you succeed in satisfying your teachers with your schoolwork? A8</td>
<td>.74</td>
</tr>
<tr>
<td>How well can you motivate yourself to do schoolwork? A6</td>
<td>.71</td>
</tr>
<tr>
<td>How well can you pay attention during class? A4</td>
<td>.70</td>
</tr>
<tr>
<td>How well can you study for a test? A2</td>
<td>.62</td>
</tr>
<tr>
<td>How well can you study when there are other interesting things to do? A1</td>
<td>.61</td>
</tr>
<tr>
<td>How well can you remember information given in class? A5</td>
<td>.59</td>
</tr>
<tr>
<td><strong>Social self-efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>How well can you work in a group? S8</td>
<td>.76</td>
</tr>
<tr>
<td>How well can you cooperate with your classmates? S4</td>
<td>.74</td>
</tr>
<tr>
<td>How well can you become friends with other people? S2</td>
<td>.73</td>
</tr>
<tr>
<td>How well can you succeed in staying friends with other people? S7</td>
<td>.69</td>
</tr>
<tr>
<td>How well can you participate in class activities? S6</td>
<td>.66</td>
</tr>
<tr>
<td>How well can you participate in class discussions? S5</td>
<td>.55</td>
</tr>
<tr>
<td>How well can you have a chat with an unfamiliar person of your age? S3</td>
<td>.53</td>
</tr>
<tr>
<td>How well can you express your opinions when your classmates disagree with you? S1</td>
<td>.39</td>
</tr>
<tr>
<td><strong>Help seeking self-efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>How well can you get the information you need from adults? Hs2</td>
<td>.35</td>
</tr>
<tr>
<td>How well can you get school staff to help you, when you have a problem at school? H1</td>
<td>.31</td>
</tr>
<tr>
<td>How well can you get school staff to help you, when you have a problem at school? Hs4</td>
<td>.37</td>
</tr>
<tr>
<td>How well can you get teachers to help you when you get stuck on schoolwork? Hs3</td>
<td>.38</td>
</tr>
</tbody>
</table>
Table 2. Principal Components Analysis of Pre-Programme Self-Efficacy Items with Varimax Rotation (2 components specified)

<table>
<thead>
<tr>
<th>Components</th>
<th>Items</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Academic self-efficacy</strong></td>
<td>How well can you finish school assignments? A7</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you succeed in finishing all your homework? A3</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you succeed in satisfying your teachers with your schoolwork? A8</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you motivate yourself to do schoolwork? A6</td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you pay attention during class? A4</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you study for a test? A2</td>
<td>.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you remember information given in class? A5</td>
<td>.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How well can you study when there are other interesting things to do? A1</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td><strong>Help seeking self-efficacy</strong></td>
<td>How well can you get teachers to help you when you get stuck on schoolwork? Hs3</td>
<td>.52</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td>How well can you get the information you need from adults? Hs2</td>
<td>.52</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>How well can you get school staff to help you, when you have a problem at school? Hs4</td>
<td>.51</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td>How well can you get adults to help you with a problem? Hs1</td>
<td>.47</td>
<td>.42</td>
</tr>
<tr>
<td><strong>Social self-efficacy</strong></td>
<td>How well can you work in a group? S8</td>
<td></td>
<td>.76</td>
</tr>
<tr>
<td></td>
<td>How well can you become friends with other people? S2</td>
<td></td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>How well can you cooperate with your classmates? S4</td>
<td></td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td>How well can you participate in class activities? S6</td>
<td></td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>How well can you participate in class discussions? S5</td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>How well can you have a chat with an unfamiliar person of your age? S3</td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>How well can you succeed in staying friends with other people? S7</td>
<td></td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>How well can you express your opinions when your classmates disagree with you? S1</td>
<td></td>
<td>.56</td>
</tr>
</tbody>
</table>

* Loadings < .30 suppressed
Table 3. Principal Components Analysis of Pre-Programme Academic and Social Self-Efficacy Items with Varimax Rotation (2 components specified)

<table>
<thead>
<tr>
<th>Items</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Academic self-efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>How well can you succeed in finishing all your homework? A3</td>
<td>.78</td>
</tr>
<tr>
<td>How well can you finish school assignments? A7</td>
<td>.78</td>
</tr>
<tr>
<td>How well can you succeed in satisfying your teachers with your schoolwork? A8</td>
<td>.76</td>
</tr>
<tr>
<td>How well can you motivate yourself to do schoolwork? A6</td>
<td>.73</td>
</tr>
<tr>
<td>How well can you pay attention during class? A4</td>
<td>.72</td>
</tr>
<tr>
<td>How well can you study for a test? A2</td>
<td>.67</td>
</tr>
<tr>
<td>How well can you remember information given in class? A5</td>
<td>.64</td>
</tr>
<tr>
<td>How well can you study when there are other interesting things to do? A1</td>
<td>.63</td>
</tr>
<tr>
<td><strong>Social self-efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>How well can you work in a group? S8</td>
<td>.77</td>
</tr>
<tr>
<td>How well can you become friends with other people? S2</td>
<td>.73</td>
</tr>
<tr>
<td>How well can you cooperate with your classmates? S4</td>
<td>.70</td>
</tr>
<tr>
<td>How well can you participate in class activities? S6</td>
<td>.70</td>
</tr>
<tr>
<td>How well can you participate in class discussions? S5</td>
<td>.70</td>
</tr>
<tr>
<td>How well can you have a chat with an unfamiliar person of your age? S3</td>
<td>.66</td>
</tr>
<tr>
<td>How well can you succeed in staying friends with other people? S7</td>
<td>.63</td>
</tr>
<tr>
<td>How well can you express your opinions when your classmates disagree with you? S1</td>
<td>.56</td>
</tr>
</tbody>
</table>

*Loadings < .30 suppressed*
Table 4. Principal Components Analysis of One Year Post-Programme Academic, Social, and Career Decision Self-Efficacy with Varimax Rotation (3 components specified)

<table>
<thead>
<tr>
<th>Components</th>
<th>Items</th>
<th>Items</th>
<th>Items</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>How well can you study for a test? A2</td>
<td>How well can you study for a test? A2</td>
<td>How well can you study for a test? A2</td>
<td>.79</td>
</tr>
<tr>
<td>2</td>
<td>How well can you motivate yourself to do schoolwork? A6</td>
<td>How well can you motivate yourself to do schoolwork? A6</td>
<td>How well can you motivate yourself to do schoolwork? A6</td>
<td>.77</td>
</tr>
<tr>
<td>3</td>
<td>How well can you succeed in finishing all your homework? A3</td>
<td>How well can you succeed in finishing all your homework? A3</td>
<td>How well can you succeed in finishing all your homework? A3</td>
<td>.77</td>
</tr>
<tr>
<td>1</td>
<td>How well can you finish school assignments? A7</td>
<td>How well can you finish school assignments? A7</td>
<td>How well can you finish school assignments? A7</td>
<td>.74</td>
</tr>
<tr>
<td>2</td>
<td>How well can you succeed in satisfying your teachers with your schoolwork? A8</td>
<td>How well can you succeed in satisfying your teachers with your schoolwork? A8</td>
<td>How well can you succeed in satisfying your teachers with your schoolwork? A8</td>
<td>.73</td>
</tr>
<tr>
<td>3</td>
<td>How well can you study when there are other interesting things to do? A1</td>
<td>How well can you study when there are other interesting things to do? A1</td>
<td>How well can you study when there are other interesting things to do? A1</td>
<td>.70</td>
</tr>
<tr>
<td>1</td>
<td>How well can you pay attention during class? A4</td>
<td>How well can you pay attention during class? A4</td>
<td>How well can you pay attention during class? A4</td>
<td>.69</td>
</tr>
<tr>
<td>2</td>
<td>How well can you remember information given in class? A5</td>
<td>How well can you remember information given in class? A5</td>
<td>How well can you remember information given in class? A5</td>
<td>.57 .34</td>
</tr>
<tr>
<td>1</td>
<td>How well can you work out what job would be best for you? Cd1</td>
<td>How well can you work out what job would be best for you? Cd1</td>
<td>How well can you work out what job would be best for you? Cd1</td>
<td>.81</td>
</tr>
<tr>
<td>2</td>
<td>How well can you decide what you are looking for in a job Cd2</td>
<td>How well can you decide what you are looking for in a job Cd2</td>
<td>How well can you decide what you are looking for in a job Cd2</td>
<td>.80</td>
</tr>
<tr>
<td>3</td>
<td>How well can you choose a further education/training course or a job that will fit your interests? Cd6</td>
<td>How well can you choose a further education/training course or a job that will fit your interests? Cd6</td>
<td>How well can you choose a further education/training course or a job that will fit your interests? Cd6</td>
<td>.79</td>
</tr>
<tr>
<td>1</td>
<td>How well can you identify employers and further education/training institutes relevant to your career path? Cd8</td>
<td>How well can you identify employers and further education/training institutes relevant to your career path? Cd8</td>
<td>How well can you identify employers and further education/training institutes relevant to your career path? Cd8</td>
<td>.78</td>
</tr>
<tr>
<td>2</td>
<td>How well can you determine the steps you need to take to successfully complete your chosen further education/training course? Cd7</td>
<td>How well can you determine the steps you need to take to successfully complete your chosen further education/training course? Cd7</td>
<td>How well can you determine the steps you need to take to successfully complete your chosen further education/training course? Cd7</td>
<td>.78</td>
</tr>
<tr>
<td>3</td>
<td>How well can you find information about further education/training courses that will help you get the job you want? Cd3</td>
<td>How well can you find information about further education/training courses that will help you get the job you want? Cd3</td>
<td>How well can you find information about further education/training courses that will help you get the job you want? Cd3</td>
<td>.75</td>
</tr>
<tr>
<td>1</td>
<td>How well can you prepare a good resume/CV? Cd4</td>
<td>How well can you prepare a good resume/CV? Cd4</td>
<td>How well can you prepare a good resume/CV? Cd4</td>
<td>.51</td>
</tr>
<tr>
<td>2</td>
<td>How well can you perform well in a job interview? Cd5</td>
<td>How well can you perform well in a job interview? Cd5</td>
<td>How well can you perform well in a job interview? Cd5</td>
<td>.51 .41</td>
</tr>
<tr>
<td>1</td>
<td>How well can you work in a group? S8</td>
<td>How well can you work in a group? S8</td>
<td>How well can you work in a group? S8</td>
<td>.74</td>
</tr>
<tr>
<td>2</td>
<td>How well can you have a chat with an unfamiliar person of your age? S3</td>
<td>How well can you have a chat with an unfamiliar person of your age? S3</td>
<td>How well can you have a chat with an unfamiliar person of your age? S3</td>
<td>.73</td>
</tr>
<tr>
<td>3</td>
<td>How well can you become friends with other people? S2</td>
<td>How well can you become friends with other people? S2</td>
<td>How well can you become friends with other people? S2</td>
<td>.72</td>
</tr>
<tr>
<td>1</td>
<td>How well can you participate in class discussions? S5</td>
<td>How well can you participate in class discussions? S5</td>
<td>How well can you participate in class discussions? S5</td>
<td>.71</td>
</tr>
<tr>
<td>2</td>
<td>How well can you express your opinions when your classmates disagree with you? S1</td>
<td>How well can you express your opinions when your classmates disagree with you? S1</td>
<td>How well can you express your opinions when your classmates disagree with you? S1</td>
<td>.69</td>
</tr>
<tr>
<td>3</td>
<td>How well can you cooperate with your classmates? S4</td>
<td>How well can you cooperate with your classmates? S4</td>
<td>How well can you cooperate with your classmates? S4</td>
<td>.68</td>
</tr>
<tr>
<td>1</td>
<td>How well can you participate in class activities? S6</td>
<td>How well can you participate in class activities? S6</td>
<td>How well can you participate in class activities? S6</td>
<td>.38 .67</td>
</tr>
<tr>
<td>2</td>
<td>How well can you succeed in staying friends with other people? S7</td>
<td>How well can you succeed in staying friends with other people? S7</td>
<td>How well can you succeed in staying friends with other people? S7</td>
<td>.58</td>
</tr>
</tbody>
</table>

*Loadings < .30 suppressed
Table 5. Sample sizes per Programme across All Outcome Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Self-Efficacy (Time 1)</td>
<td>21.8</td>
<td>22.0</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 1)</td>
<td>21.8</td>
<td>22.0</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 2)</td>
<td>17.8</td>
<td>19.0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 2)</td>
<td>17.8</td>
<td>19.0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 3)</td>
<td>9.7</td>
<td>9.5</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 3)</td>
<td>11.9</td>
<td>12.0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Career Decision Self-Efficacy (Time 3)</td>
<td>11.9</td>
<td>12.0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>NCEA Level 1 Credits (Time 3)</td>
<td>15.4</td>
<td>19.0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>NCEA Level 1 Status (Time 3)</td>
<td>15.4</td>
<td>19.0</td>
<td>0</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 6. Frequency of Participants Associated with each Outcome Variable at each Relevant Time Point for All Programmes Combined

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pre-Programme (Time 1)</th>
<th>Post-Programme (Time 2)</th>
<th>One Year Post-Programme (Time 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PK Control</td>
<td>PK Control</td>
<td>PK Control</td>
</tr>
<tr>
<td>Academic Self-Efficacy</td>
<td>514 576</td>
<td>471 420</td>
<td>251 236</td>
</tr>
<tr>
<td>Social Self-Efficacy</td>
<td>514 576</td>
<td>471 420</td>
<td>309 285</td>
</tr>
<tr>
<td>Career Decision Self-Efficacy</td>
<td>- -</td>
<td>- -</td>
<td>309 284</td>
</tr>
<tr>
<td>NCEA Level 1 Credits</td>
<td>- -</td>
<td>- -</td>
<td>390 401</td>
</tr>
<tr>
<td>NCEA Level 1 Status</td>
<td>- -</td>
<td>- -</td>
<td>389 401</td>
</tr>
<tr>
<td>Variables</td>
<td>Pre-Programme (Time 1)</td>
<td>Post-Programme (Time 2)</td>
<td>One Year Post-Programme (Time 3)</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------</td>
<td>-------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td>PK</td>
<td>Control</td>
<td>PK</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td>237 (46.1%)</td>
<td>276 (48.1%)</td>
<td>216 (45.9%)</td>
</tr>
<tr>
<td>European</td>
<td>395 (76.8%)</td>
<td>398 (69.1%)</td>
<td>368 (78.1%)</td>
</tr>
<tr>
<td>Māori</td>
<td>119 (23.2%)</td>
<td>145 (25.2%)</td>
<td>108 (22.9%)</td>
</tr>
<tr>
<td>Pacific Peoples</td>
<td>69 (13.4%)</td>
<td>98 (17.0%)</td>
<td>59 (12.5%)</td>
</tr>
<tr>
<td>Asian</td>
<td>23 (4.5%)</td>
<td>38 (6.6%)</td>
<td>21 (4.5%)</td>
</tr>
</tbody>
</table>

Table 7. Frequency of Participants Associated with the Gender and Ethnicity Variables at each Relevant Time Point for All Programmes Combined
**Results of Baseline Differences between Project K Graduates and Non-Graduates**

There was a significantly greater proportion of participants identifying with a European ethnicity in the Project K graduate group \( (\chi^2 (1, N = 585) = 12.73, p. = .000) \) whereas the proportion of participants identifying with a Pacific Peoples ethnicity in the sample of non-graduates was higher than those in the graduate sample \( (\chi^2 (1, N = 585) = 7.04, p. = .008) \).

Independent samples t-tests were used to investigate differences between these two groups in their academic and social self-efficacy scores at baseline and these results revealed that the participants who did complete the programme had significantly higher pre-programme social self-efficacy scores than those who did not \( (t (559) = -2.27, p. = .02) \).

**Results of Project K and Control Group Attrition on Baseline Equivalence**

Baseline self-efficacy comparisons between those who had missing self-efficacy data at the immediate post-programme time point and those with data were made using independent samples t-tests but participants did not differ on academic or social self-efficacy just prior to the programme start date.

Chi-square analyses were conducted to assess differences between these groups the proportions of each gender and ethnicity. The only significant differences obtained between those with and without post-programme self-efficacy data were for the proportion of European \( (\chi^2 (1, N = 1076) = 11.34, p. < .001) \), Māori \( (\chi^2 (1, N = 1076) = 5.21, p. < .02) \), and Pacific Peoples \( (\chi^2 (1, N = 1076) = 4.50, p. < .03) \). To qualify these differences, a larger proportion of Europeans were found in the group with complete post-programme data, while higher numbers of Māori and Pacific Peoples did not have responses for the two outcome variables relative to those who did.

T-tests revealed significant differences on pre-programme social self-efficacy \( (t (1088) = -2.52, p. = .01) \) in comparisons of those who dropped out of the evaluation at the one year post-programme follow up to those who did participate. Chi-square analyses indicated that the two groups were equivalent only with regards to the gender distribution. Larger proportions of Europeans \( (\chi^2 (1, N = 1076) = 18.46, p. < .000) \) and Asian participants \( (\chi^2 (1, N = 1076) = 12.48, p. < .000) \), as
well as those identifying with an “Other” ethnicity provided responses ($\chi^2 (1, N = 1076) = 7.67, p < .01$) to the social self-efficacy post-programme in comparison to the group who did not, whereas larger proportions of Māori ($\chi^2 (1, N = 1076) = 11.87, p < .001$) and Pacific Peoples ($\chi^2 (1, N = 1076) = 4.93, p < .03$) did not respond to this subscale.

Table 8. Results from Exploration of Normality Assumptions for Full Participant Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Test of Normality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>S</td>
<td>S_SE</td>
</tr>
<tr>
<td>Level 1 Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 1)</td>
<td>1092</td>
<td>.12</td>
<td>.07</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 2)</td>
<td>891</td>
<td>-.06</td>
<td>.08</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 3)</td>
<td>487</td>
<td>-.18</td>
<td>.11</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 1)</td>
<td>1092</td>
<td>-.39</td>
<td>.07</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 2)</td>
<td>891</td>
<td>-.38</td>
<td>.08</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 3)</td>
<td>594</td>
<td>.48</td>
<td>.10</td>
</tr>
<tr>
<td>Career-Decision Self-Efficacy (Time 3)</td>
<td>593</td>
<td>-.06</td>
<td>.10</td>
</tr>
<tr>
<td>NCEA Level-1 Credits (Time 3)</td>
<td>791</td>
<td>.11</td>
<td>.09</td>
</tr>
<tr>
<td>Level 2 Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decile Rating</td>
<td>50</td>
<td>-.09</td>
<td>.34</td>
</tr>
</tbody>
</table>

Notes. *** p < .001  
** p < .01  
* p < .05  
*p < .10
<table>
<thead>
<tr>
<th>Variable</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Test of Normality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>$S$</td>
<td>$S_{SE}$</td>
</tr>
<tr>
<td><strong>Level 1 Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 1)</td>
<td>516</td>
<td>.14</td>
<td>.11</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 2)</td>
<td>471</td>
<td>-.14</td>
<td>.11</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 3)</td>
<td>251</td>
<td>-.25</td>
<td>.15</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 1)</td>
<td>516</td>
<td>-.36</td>
<td>.11</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 2)</td>
<td>471</td>
<td>-.42</td>
<td>.11</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 3)</td>
<td>309</td>
<td>-.47</td>
<td>.14</td>
</tr>
<tr>
<td>Career-Decision Self-Efficacy (Time 3)</td>
<td>309</td>
<td>-.17</td>
<td>.14</td>
</tr>
<tr>
<td>NCEA Level-1 Credits (Time 3)</td>
<td>390</td>
<td>.01</td>
<td>.12</td>
</tr>
</tbody>
</table>

**Notes.*** $p < .001$  
**$p < .01$ 
* $p < 0.05$  
$p < .10$
Table 10. Results from Exploration of Normality Assumptions for Control Participant Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Test of Normality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>S</td>
<td>S_SE</td>
</tr>
<tr>
<td>Level 1 Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 1)</td>
<td>516</td>
<td>.11</td>
<td>.10</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 2)</td>
<td>471</td>
<td>-.01</td>
<td>.12</td>
</tr>
<tr>
<td>Academic Self-Efficacy (Time 3)</td>
<td>251</td>
<td>-.11</td>
<td>.16</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 1)</td>
<td>516</td>
<td>-.42</td>
<td>.10</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 2)</td>
<td>471</td>
<td>-.25</td>
<td>.12</td>
</tr>
<tr>
<td>Social Self-Efficacy (Time 3)</td>
<td>309</td>
<td>-.40</td>
<td>.14</td>
</tr>
<tr>
<td>Career-Decision Self-Efficacy (Time 3)</td>
<td>309</td>
<td>.03</td>
<td>.15</td>
</tr>
<tr>
<td>NCEA Level-1 Credits (Time 3)</td>
<td>390</td>
<td>.21</td>
<td>.12</td>
</tr>
</tbody>
</table>

*** p < .001
** p < .01
* p < 0.05
*p < .10
Appendix F: Example Outputs from Stakeholder Empowering Processes
WHAT WOULD YOU LIKE TO KNOW ABOUT PROJECT K?

Over the past few years, lots of information has been collected to help us assess Project K; now is the time to get the most we can out of the results. Below is a list of questions. As explained in the information sheet, I (Kelsey Deane) am trying to get an idea of which questions about Project K are the most important to different stakeholders. Please think carefully about which questions will provide answers that will be the most useful to you and/or your community.

You have the opportunity to use 5 votes to select your priority questions.

You may divide the votes any way you like (for example, if you feel very strongly that one question should be answered you may want to place more than one of your votes OR all of your votes on that question, OR perhaps you may want to select 5 different questions of equal importance).

NOTE: I will not be able to answer some of the questions presented below within the timeframe and resources available during my PhD research, but I’ve included them so that FYD can get a thorough picture of what everyone’s different evaluation needs are.

Once I obtain the votes, I will let FYD and Project K licensees know the results and the plan for how the most burning questions will be addressed.

If you would like a research summary of the results obtained please indicate your email address OR name and postal address below and separate this sheet from your votes (to ensure your votes remain anonymous):

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Please indicate your choices clearly. You may only use a maximum of 5 votes, but each number below represents one possible voting choice (i.e. if there is more than one question associated with a specific number that group of questions represents just ONE choice).

1. What is Project K in concept (what is it supposed to be doing) & in practice (what is it actually doing)? Are Project K’s operations consistent with Project K’s conceptual theory?

2. To what extent was each Project K objective achieved (based on comparing the results of Project K and Control students on the measures collected (i.e. self-efficacy, health & lifestyle behaviours, social competence, NCEA results, etc.)? Were these outcomes maintained (at one-year and three-years post-programme)?

3. What is Project K’s effect on outcomes (this refers to outcomes in addition to those that were measured --with standardized surveys, etc.)?

4. What are the long-term (broad) impacts of Project K?

5. Is Project K working? Should it be revised?

6. What changes in the Project K’s design or implementation might produce better outcomes? For which types of individuals will these changes have the greatest impact? For which types of individuals will these changes have little impact?

7. Why does Project K work (or if it doesn’t, why not)? How does Project K produce outcomes? What factors may cause or contribute to the obtained outcomes? What were the most important reasons for Project K’s success or failure? What Project K features are essential for successful replication?

8. What are Project K’s strengths and shortfalls? Did Project K have any unintended negative or positive side effects?

9. What are Project K’s most important unresolved issues?

10. What facilities, personnel, materials and equipment are needed for improving Project K?

11. For which types of young people does Project K work best? Why?

12. What are the strengths and shortfalls within the specific Project K components (i.e. Wilderness Adventure, Community Challenge, and Mentoring)?

13. For which types of individuals do specific components (Wilderness Adventure, Community Challenge, Mentoring) of Project K work best? Why?

14. In which contexts (within which regions – rural, suburban, urban; decile ranks; licensee/community characteristics) does Project K work best? Why?
15. What sectors (regions, decile rank levels, licensee programmes, ethnicities, gender) within Project K are performing best and poorest? How did each sector of Project K perform in relation to each concept measured (self-efficacy, health & lifestyle behaviours, social competence, NCEA, etc)?

16. Why did different sectors of Project K (regions, decile rank levels, licensee programmes, ethnicities, gender) perform the way they did on each concept measured (self-efficacy, health & lifestyle behaviours, social competence, NCEA, etc)?

17. What are the strengths & shortfalls within specific Project K licensee programmes? (This refers to strengths and shortfalls that can be identified by performance on the measures collected but also strengths and opportunities for improvement that are additional to these).

18. Is Project K’s present performance better than past performances? Are specific Project K licensee programmes’ present performances better than past performances (in relation to performance on the concepts that were measured)?

19. Did Project K meet all the participating students’ needs? What participating students’ needs should be addressed that are not currently being addressed?

20. Is Project K reaching the right students (i.e. those with low self-efficacy)?

21. Did Project K meet any community’s needs? Which community needs did Project K meet? How much can Project K be adapted to meet various cultural/community needs?

22. Did Project K meet any of the participating students’ families’ needs? Which family needs did Project K meet?

23. How do various stakeholders (including experts) value Project K (How do their opinions of Project K differ)?

24. What interesting stories related to Project K are out there?

25. How has Project K evolved over time? How should Project K evolve in the future?

26. Is Project K worthy of continuation and/or further distribution?

27. Is Project K sustainable & transportable?

28. Did Project K effectively discharge its responsibilities?

29. What were the costs per student, per year, etc? What are the costs vs. the benefits of operating Project K? Is Project K worth the required investment? Is there any way of significantly reducing the cost of delivering Project K while still retaining safety and effectiveness?
30. What alternatives to Project K are available to address the students'/families'/communities' needs & what are their comparative merits? Is Project K superior to critical competitors?

31. How can we ensure Project K is delivered to greater numbers?

32. Would Project K work on other client groups? If yes, which ones?

33. Can you think of any questions that were not represented in the above list? Please list them below and clearly indicate if you would like to vote for one or more of these.

To identify if there are any differences in responses based on different stakeholder roles within the Project K organisation please answer the following (optional):

I am a:

- Foundation for Youth Development National Support Office Staff Member
- Project K Licensee Staff Member
- Foundation for Youth Development Trustee
- Project K Trustee
- Programme Sponsor/Funder
- Mentor
- Other Volunteer
- Project K participant (past or current)
- Project K participant caregiver
- School staff member
- Other ____________________
Please return this completed questionnaire and the separate sheet with your contact details to a Project K staff member otherwise you are welcome to email your votes directly to k.deane@auckland.ac.nz or contact Kelsey by telephone at 373 7599 ext. 84204. Thank You!

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE for a period of three years from 18/02/2008 to 09/04/2012. Reference 2008/009.
<table>
<thead>
<tr>
<th>Broad Theme</th>
<th>Subcategory</th>
<th>Relevant Programme Development Opportunity</th>
<th>Example Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greater Integration</td>
<td>Integration across the 3 programme components/within programme</td>
<td>Engagement in Community Challenge</td>
<td>• Provider staying linked with PK students</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participant Reintegration</td>
<td>• Continued provider support during mentor phase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Health &amp; Fitness Outcomes</td>
<td>• Mentor integration in WA to promote empathy</td>
</tr>
<tr>
<td></td>
<td>With young person’s external support networks</td>
<td>All Four</td>
<td>• Link health and fitness outcomes with mentor/mentor activities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Group activities ongoing through year</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Learnings + goals from Wilderness being outlined with students at beginning of CC – daily goals/reflection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Link CC with FYD sponsors, Community Engagement programmes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Need to use the relationships developed in the wilderness more. Eg. Some planned days away with mentors and students – and continued therapeutic use of wilderness staff and students</td>
</tr>
<tr>
<td></td>
<td>With Maori culture</td>
<td>Cultural Consideration</td>
<td>• More support from schools-work closely with liaison</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Consider reintegration at end of 14 months (as well)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Introduce to community during WA – stay with and help members of the public, make students aware when these people help us.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Whanau involvement in Community Challenge (i.e. family ropes course day)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Assembly presentation – when return for school staff, peers, whānau, mentors – all these groups at once.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Teacher presentations before and after WA.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Participants encouraging parents, bros &amp; sis – good food at home</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Significance of local area/marae explained by local Kaumatua</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Finding young Maori role models for Community Challenge</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Utilising marae visits. Pouhiri experiences to bring focus &amp; extend understanding</td>
</tr>
<tr>
<td>Broad Theme</td>
<td>Subcategory</td>
<td>Relevant Programme Development Opportunity</td>
<td>Example Ideas</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Educational sessions/workshops</td>
<td>For external support networks</td>
<td>Participant Reintegration Health &amp; Fitness Outcomes</td>
<td>• Teacher presentations before and after WA.                                                                                                     • Health &amp; fitness education for community/whanau as part of the community challenge</td>
</tr>
<tr>
<td></td>
<td>For participants</td>
<td>Participant Reintegration Health &amp; Fitness Outcomes</td>
<td>• Possibly develop a workshop for whanau re: reintegration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Education on diet &amp; fitness                                                                                                                                  • Pre-wilderness dietary &amp; fitness workshops – maybe at survey stage</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Substance workshops                                                                                                                                        • Afternoon tutorials after CC</td>
</tr>
<tr>
<td>Meaningful Activities &amp; Lessons</td>
<td>Student Ownership</td>
<td>Engagement in Community Challenge</td>
<td>• Punishments/rewards outlined at beginning and being identified by students                                                                                                                                  • Ownership – it’s not a guided tour</td>
</tr>
<tr>
<td></td>
<td>Relevant/Fun</td>
<td>Engagement in Community Challenge</td>
<td>• Relevant CC days                                                                                                                                                                                       • Helping people better than fence painting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Making learning fun                                                                                                                                                                                      • Students researched what the community wanted from them</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Equality for all                                                                                                                                                                                         • We forget school is their 1st community</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Helping students understand “What is community?”                                                                                                                                                        • Set up CC for success: Preplan project where there is a connection for the 1st day.</td>
</tr>
<tr>
<td>Need for workforce training</td>
<td>Cultural Consideration</td>
<td></td>
<td>• Lack of easy/simple ways to upskill                                                                                                                                                                       • Mismatch between client/student cultural identity &amp; providers</td>
</tr>
<tr>
<td></td>
<td>Multiculturalism</td>
<td></td>
<td>• If the treaty is one of our guiding principles then we have to have an infrastructure which supports our workforce with taha Maori such as appointment of kuia/kaumatua/cultural advisors – NCYW code of conduct</td>
</tr>
<tr>
<td>Broader concept for cultural awareness</td>
<td>Multidimensional</td>
<td>Cultural Consideration</td>
<td>• It’s not just Maori, so many cultures, need time to learn                                                                                                                                             • Equality for all</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Also learning about other cultures                                                                                                                                                                        • Experiencing different cultures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• How many PK students are Asian and Pakeha – How is it perceived?                                                                                                                                         • Culture is not just about ethnicity. Culture has many forms. Eg. Youth culture</td>
</tr>
<tr>
<td></td>
<td>Multidimensional</td>
<td>Cultural Consideration</td>
<td>• Developing pride in community</td>
</tr>
<tr>
<td>Broad Theme</td>
<td>Subcategory</td>
<td>Relevant Programme Development Opportunity</td>
<td>Example Ideas</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>Engagement in Community Challenge Participant Reintegration Health &amp; Fitness Outcomes</td>
<td></td>
<td>• Reward all students. Have eye on WA. Revisit values as reward • Acknowledgement • Reward for participation (Kayaking) depending on attendance</td>
</tr>
<tr>
<td>Other specific suggestions</td>
<td>Cultural Consideration Health &amp; Fitness Outcomes</td>
<td></td>
<td>• Blessing on WA departure &amp; return + at meal times • Break meals down into per person &amp; compare to McD’s. • Eating &amp; Fitness diary • Introduction to clubs • Student meal planning &amp; purchasing • During CC introduce to new sports/gym/clubs • Visit youth health clinic – sexual health/drugs/alcohol in CC. • Model healthy eating &amp; lifestyle • Goals during mentoring for fitness</td>
</tr>
<tr>
<td>Engagement in the Community Challenge</td>
<td></td>
<td></td>
<td>• Amazing race – SPCA, AA, Police, IRD, Bank, Polytech, etc. • Don’t put through complication. Simple &amp; effective</td>
</tr>
</tbody>
</table>

Notes. Example ideas are in stakeholders own words as written on worksheets.
References


Retrieved from
x?direct=true&db=aph&AN=3078315&site=ehost-live&scope=site


Lakin, R., & Mahoney, A. (2006). Empowering youth to change their world: Identifying key components of a community service program to promote positive development. *Journal of School Psychology, 44*(6), 513-531. doi: http://dx.doi.org/10.1016/j.jsp.2006.06.001


Roth, J., Qiang, X., Marban, S. L., Redelt, H., & Lowell, B. C. (2004). The obesity pandemic: Where have we been and where are we going? *Obesity, 12*(90002), 88S-101S.


emotional, and personality development (6th ed.) (pp. 933-1002). Hoboken, NJ: John Wiley & Sons Inc; US.


