

<http://researchspace.auckland.ac.nz>

**University of Auckland Research Repository, ResearchSpace**

### **Copyright Statement**

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

### **General copyright and disclaimer**

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](#) and [Deposit Licence](#).

Investigating the implementation and impact of participation in a health  
promotion programme on the social attitudes and behaviours of children in  
New Zealand schools

Rachel Allyson Williams

A thesis submitted in fulfillment of the requirements for the degree of  
Doctor of Philosophy (PhD) in Education, The University of Auckland, 2015

## Abstract

The social health and wellbeing of vulnerable children in New Zealand, particularly those from low-decile neighbourhoods, is not ideal. The New Zealand government and the New Zealand public have called for educators to redress this issue through the delivery of a comprehensive school-based health education programme. To date, however, little is known about the implementation or outcomes of participation in school-based health education classes, most notably under low-decile, highly transient school conditions. This PhD thesis, therefore, investigated the implementation of the traditional health education curriculum and, subsequently, evaluated the impact of participation in an externally provided health promotion programme on the social attitudes and behaviours of children attending low-decile primary and intermediate schools around the country.

The results of the traditional health education study indicated that teachers ( $n = 198$ ) were delivering social health education classes infrequently, if at all, and that many schools ( $n = 61$ ) were using external providers for social health education provision. The implementation of one externally provided programme, *Kiwi Can*, was then explored and a stratified sample of Year 4-6 students ( $n = 319$ ) reported that *Kiwi Can* was being delivered with fidelity, adherence to participant responsiveness, and to a high standard (quality) in all participating regions. It was followed by a longitudinal, quasi-experimental outcome evaluation in which pre-post survey data were gathered from control ( $n = 456$ ), *new Kiwi Can* ( $n = 464$ ), and *experienced Kiwi Can* ( $n = 299$ ) school students. Significant differences in self-reported prosocial attitudes and behaviours were explored as a function of school condition and transience level using confirmatory factor analysis and structural equation modelling techniques. A hierarchical, autoregressive structural model most closely fit the data. Further, subscale means for measures of positive social attitudes and behaviours (PYD) and classroom climate (CC) decreased for all three groups over time. They decreased significantly more for new *Kiwi Can* students than for experienced *Kiwi Can* students under highly transient school conditions.

These results suggest that sustained participation in Kiwi Can may have attenuated the negative effects that a highly transient environment had on the social attitudes and behaviours of vulnerable children who were not themselves transient. This finding has significant implications for both the design and delivery of school-based, social health promotion programming for children.

## **Dedication**

*For Aronui James Punehu Williams*

## Acknowledgments

Alone we can do so little; together we can do so much.

~Helen Keller

This PhD has been a true labour of love. It has involved the active participation of so many individuals, each of whom I feel fortunate to have met, and in many cases, befriended over the past three and a half years. Two of the most pivotal of these are my thesis supervisors – Associate Professor Gavin Brown and Dr. Pat Bullen. Thank you, Gavin and Pat, for sharing your expertise and your time with me. I felt extremely well supported throughout the entire process and could not have asked for more compassionate or capable scholars to guide me on this journey. You have both inspired and motivated me. Thank you very much.

Thank you, as well, to the wonderful research and evaluation team at the Foundation for Youth Development (FYD). In particular, I would like to thank Ms. Julie Moore for her energy, support, commitment and willingness to embrace the research project from the start. Without her involvement the entire second half of this PhD project would not have been possible. On that note, I'd also like to thank Ms. Victoria Hood and curriculum and pedagogy team at FYD as well as all of the Kiwi Can programme co-ordinators and leaders for their valuable participation. It was a pleasure getting to know each of you and to working with you to design and collect data from across New Zealand. I have many lovely memories of meaningful conversations held over cups of tea with 'Kiwi Cannors' from as far north as Whangaroa and as far south as Bluff. It has been a pleasure sharing this experience with you.

It has also been a true pleasure to work with a number of capable support staff. I'd like to thank Mr. Andrew Lavery for his work transcribing interviews, Dr. Kane Meissel, and Mr. Cole Armstrong, for their statistical expertise and guidance in working through the structural equation modelling procedures. I would also like to thank the University of Auckland International Doctoral Scholarship team for the incredible financial support that they have provided over the past three and a half years. It made the entire thesis possible.

This PhD would also not have been possible without the active involvement of each and every one of my participants – primary school health education teachers, health education curriculum and positive youth development theorists, Kiwi Can and control school students, parents, teachers, and principals from schools across the country. You gave a huge amount of time and energy to participate in this research project and I am forever grateful. Thank you.

At this time I would also like to acknowledge and give particular thanks to the inspiring women that I work alongside every day – Katharina Limbach, Nike Franke, Lisa Darragh, Jennifer Tatabe, Sarah Hart, and Melanie Palmer. I feel so lucky to have such a wonderful group of friends at the university who have really bonded together to show each other care and support during this experience. I also feel incredibly lucky that Katharina, Nike, Lisa, and Sarah can each relate to the unique set of challenges that one faces when completing a doctorate while also being a mother. There is no doubt we now share a very special bond that I hope will last for years to come.

Finally, I would like to give huge thanks to my incredible family, both here in Auckland overseas. Thank you, Aunt Kelsey, for your quick wit and listening ear. Our weekly Skype dates were instrumental in helping me maintain a sense of calm and confidence that the PhD would come together (eventually)! Leigh - thank you, for being such a generous and loving sister. Andy, thank you for your unwavering support, your kindness, and for the listening ear that you provide. Lastly, mom and dad, I don't think that I will ever be able to thank you enough for everything that you have done to show me infinite care, consideration, and support. This PhD is as much mine as it is yours. It represents a conviction and commitment to exploring ways in which we can contribute to the world around us, to promote positive child health, and to enhance wellbeing for our most vulnerable population. It represents a value system and a way of being that you have role modelled for me from the beginning. I consider myself extremely lucky to have you as parents and hope that my son will, one day, be half as proud of me as I am of you. And, on that note, I would also like to thank my beautiful son, Aronui. I love you so much, Boo, and I am grateful that we have had the opportunity to share this experience together. There is no doubt in my mind that I could not have achieved this without you serving as my inspiration. Thank you.

## Table of Contents

Abstract	ii
Dedication	iv
Acknowledgments	v
List of Tables	xii
List of Figures	xv
<b>CHAPTER 1: INTRODUCTION</b>	
1.1    Child social health	1
1.2    The situation in New Zealand	2
1.3    Significance of research	6
1.4    Background to research	7
1.5    Assumptions	8
1.6    Structure of thesis	8
<b>CHAPTER 2: LITERATURE REVIEW</b>	
2.1    Introduction	11
2.2    Child poverty in New Zealand	11
2.3    Transience	13
2.3.1    Transience in New Zealand	14
2.3.2    The education system in New Zealand	15
2.3.3    Transience in New Zealand schools	16
2.3.3.1    Impact of transience on classroom climate	17
2.3.3.2    Measuring classroom and school climate	18
2.4    Poverty, transience and social health	20
2.4.1    Poverty and social health	20
2.4.2    Transience and social health	21
2.4.3    Gender, adolescence, and social health	22
2.4.4    Non-transient children in transient schools	23
2.5    The history of social health programming and education	23
2.5.1    Disease prevention approach	24
2.5.2    Health promotion	26
2.6    Positive youth development theory	28
2.6.1    Developmental assets framework	29
2.6.2    Lerner's 5C's theory of positive youth development	33

2.7	Positive youth development programmes	36
2.8	Health education curriculum in New Zealand	39
2.9	Research Questions	42
CHAPTER 3: THEORETICAL METHODOLOGY		
3.1	Programme evaluation theory	47
3.1.1	CIPP model: An educational evaluation model	49
3.2	Experimental design and sampling	51
3.2.1	Study design	51
3.2.2	Response scale design and analysis	53
3.2.3	Scales	54
3.2.4	Sampling	55
3.2.4.1	Sampling framework	56
3.3	Data preparation	57
3.3.1	Imputation	58
3.4	Data analysis	61
3.4.1	Model design	61
3.4.1.1	Assessing model fit	62
3.4.1.2	Model misspecification	63
3.4.1.3	Model selection	64
3.4.1.4	Invariance testing	64
3.4.1.5	Modeling change over time	65
3.4.2	Statistical data analysis procedures	67
3.4.2.1	Cluster analysis	67
3.4.2.2	Paired samples t-tests	69
3.4.2.3	Analysis of variance testing – ANOVAs	70
3.4.2.4	Post hoc procedures – ANOVAs	70
3.4.2.5	Multiple analysis of variance tests	71
3.4.2.6	Repeated measures ANOVAs	72
3.4.2.7	Multiple regression	73
3.4.2.8	Calculating effect size	75
3.5	Summary	75

## CHAPTER 4: THE IMPLEMENTATION EVALUATION OF THE HEALTH EDUCATION CURRICULUM IN NEW ZEALAND SCHOOLS

4.1	Introduction	77
4.2	Study design	78
4.2.1	Participants	78
4.2.2	Instrument	80
4.3	Data collection	82
4.4	Data analysis	82
4.5	Results	83
4.5.1	Open-ended comment data	83
4.5.2	The delivery of the health education curriculum	84
4.5.2.1	Format of health education delivery	84
4.5.2.2	Frequency of delivery of health dimensions	85
4.5.2.3	Frequency of delivery of KLAs and AOs	86
4.5.3	Teacher beliefs	91
4.5.3.1	Importance	91
4.5.3.2	Benefit	95
4.5.4	Barriers to delivery	98
4.5.4.1	Barriers and frequency of delivery in health dimensions	99
4.5.4.2	Barriers and frequency of delivery in KLAs and AOs	100
4.6	Conclusion	101
4.6.1	Implications and future directions	102

## CHAPTER 5: THE IMPLEMENTATION EVALUATION OF KIWI CAN

5.1	Introduction	103
5.2	The Kiwi Can programme	103
5.3	The implementation evaluation of Kiwi Can – Introduction	108
5.4	Study design	109
5.4.1	Sampling frame	109
5.4.2	Participants	110
5.4.3	Instrument	111
5.5	Data preparation and analysis	112
5.6	Results - Recalled values and practices (Part B)	113

5.7	Results – Kiwi Can participation (Part C)	114
5.8	Conclusion	118
CHAPTER 6: THE OUTCOME EVALUATION OF KIWI CAN		
6.1	Introduction	121
6.2	Methods – Strategy and design	123
6.2.1	Instrument	123
6.2.2	Sampling	126
6.2.3	Recruitment procedures	127
6.3	Data collection	128
6.4	Data entry and screening	129
6.4.1	Participants	130
6.5	Data analysis	131
6.6	Results – Positive social development (PYD) data	132
6.6.1	Discussion of results for PYD model	138
6.7	Modelling the classroom climate (CC) data	139
6.7.1	Discussion of results for CC model	142
6.8	The relationship between PYD and CC	142
6.9	Designing and testing the SEM	143
6.9.1	Model 1: Inter-correlated PYD and CC	144
6.9.2	Model 2: PYD to CC	148
6.9.3	Model 3: CC to PYD	151
6.9.4	Model 4: Hierarchical PYD to hierarchical CC	154
6.9.5	Model4A: Hierarchical CC to hierarchical PYD	155
6.10	Modelling change across time	158
6.11	Summary	162
6.12	Results - Mean differences in PYD and CC subscales	162
6.12.1	Baseline means – PYD subscales	162
6.12.2	Baseline means – CC subscales	164
6.12.3	Discussion of baseline PYD and CC subscale means	166
6.12.4	Change in PYD and CC subscale means over time	167
6.12.4.1	PYD by school condition	167
6.12.4.2	PYD by participant characteristics	168

6.12.4.3	CC by school condition, sex, year level, and ethnicity	168
6.12.5	Discussion of change in PYD and CC subscale mean results	168
6.13	Exploring changes in PYD and CC subscale means – Transience	171
6.13.1	Defining transience	172
6.14	Results – Baseline means (Transience)	175
6.14.1	Baseline means – PYD subscales	175
6.14.2	Baseline means – CC subscales	176
6.14.3	Discussion of baseline means	177
6.15	Results – Change in subscale means over time (Transience)	177
6.16	Results – Interactive effect of transience and school category on means	180
6.16.1	Baseline means – PYD and CC subscales	180
6.16.2	Change in subscale means – PYD and CC	182
6.16.3	Summary of results for PYD and CC subscales	185
<b>CHAPTER 7: DISCUSSION</b>		
7.1	Overview of results	186
7.2	Implications	189
7.2.1	Implications for primary school health educators and external providers	189
7.2.2	Implications for curriculum and teacher professional development	192
7.2.3	Implications for positive youth development theory	193
7.3	Successes and challenges	194
7.4	Future directions	196
7.5	Contributions	199
Appendices		201
References		266

## List of Tables

Table 1: Frequency count of demographic characteristics for health education survey participants	80
Table 2: Self-reported frequency of delivery across the four health dimensions	85
Table 3: Means, standard deviations, and output from paired samples t-tests for the KLAs in the New Zealand health education curriculum	86
Table 4: Means, standard deviations, and output from paired samples t-tests for delivery of the achievement objectives (AOs) in the New Zealand HPE curriculum	89
Table 5: Means, standard deviations, confidence intervals, and output of paired samples t-tests for importance of teaching the four dimensions of health on student wellbeing	91
Table 6: Inter-correlations for importance and frequency of delivery across the health dimensions	92
Table 7: Multiple regression output for personal beliefs about the importance of a health dimension as a predictor of delivery in health areas	93
Table 8: Inter-correlations for importance and frequency of delivery of KLAs	93
Table 9: Inter-correlations for importance and frequency of delivery of AOs	94
Table 10: Means, standard deviations, confidence intervals and output of paired samples t-tests for teaching the four dimensions of health on student wellbeing	95
Table 11: Inter-correlations for benefit and frequency of delivery across the health dimensions	96
Table 12: Inter-correlations for benefit and frequency of delivery in AOs	97
Table 13: Means, standard deviations, and paired samples t-test results for self-reported barriers by primary and intermediate school health education practitioners	98
Table 14: Multiple regression output for predictive effect of barriers on the frequency of delivery across health dimensions	100
Table 15: Multiple regression output for predictive effect of self-reported barriers and frequency of delivery across KLAs and AOs	101
Table 16: Themes and modules in the Kiwi Can programme curriculum	107
Table 17: Sampling frame for the Kiwi Can implementation evaluation	109
Table 18: Frequency count of participants by demographic variable	110

Table 19: Latent factors and items from the Kiwi Can implementation evaluation questionnaire	112
Table 20: Frequency of student responses for the learning values and social skill development data from the Kiwi Can programme questionnaire (Part B)	114
Table 21: Invariance testing for three-factor SEM by condition	116
Table 22: Mean scores for liking Kiwi Can, learning in Kiwi Can, and Kiwi Can programme by school condition, sex, year level, and ethnicity	117
Table 23: Scales, subscales, and sample items for the positive social development, PYD, section of the Kiwi Can outcome evaluation questionnaire	125
Table 24: Scales, subscales, and sample items for the classroom climate, CC, section of the Kiwi Can outcome evaluation questionnaire	126
Table 25: Sampling framework for Kiwi Can outcome evaluation	127
Table 26: Frequency counts for student participants in Kiwi Can outcome evaluation by school condition, sex, year level, and ethnicity	131
Table 27: Goodness-of-fit statistics for PYD measurement models	135
Table 28: Scale validation for PYD data at time 1 and time 2	137
Table 29: Goodness-of-fit statistics for CC measurement models	139
Table 30: Scale validation for PYD data at time 1 and time 2	142
Table 31: Results of invariance testing for final PYD and CC measurement models	143
Table 32: Goodness-of-fit statistics for PYD and CC structural equation models	145
Table 33: Covariances for inter-correlated PYD and CC structural equation Models	147
Table 34: Results of invariance testing for inter-correlated PYD-CC structural equation model	148
Table 35: Regression weights for Model 2A: PYD predicts CC	150
Table 36: Results of invariance testing for the final PYD predicts CC structural equation model	151
Table 37: Significant regression weights for Model 3A: CC predicts PYD	153
Table 38: Results of invariance testing for the final CC predicts PYD structural equation model	154
Table 39: Results of invariance testing for Model 4 and Model 4A	157
Table 40: Regression weights for hierarchical structural equation model: Model 4 and Model 4A	158

Table 41: Results of invariance testing for Model 7	159
Table 42: Standardised regression weights for time 1 and time 2 PYD and CC factors	159
Table 43: Games-Howell post hoc test results (MANOVA) for effect of school condition on PYD subscale scores at baseline	163
Table 44: Games-Howell post hoc test results (MANOVA) for effect of sex, year level, and ethnicity on PYD subscale scores at baseline	164
Table 45: Games-Howell post hoc test results (MANOVA) for effect of year level, and ethnicity on CC subscale scores at baseline	165
Table 46: One-way repeated measures analysis of variance results and effect sizes for pre-post PYD subscale means as a function of school condition	167
Table 47: The Nap, Npa, NT1, Nenroll, legit and transience scores for the participating schools	174
Table 48: Means, standard deviations, and confidence intervals for PYD subscales as a function of transience	176
Table 49: Comparison of means for CC factors at time 1 between students from low, middle, and high transience schools	177
Table 50: Mean change scores, standard deviations, and confidence intervals for PYD and CC subscales as a function of transience	178
Table 51: One-way repeated measures analysis of variance results and effect sizes for pre-post PYD and CC subscale means as a function of transience	179
Table 52: Means and standard deviations for PYD and CC subscales at baseline as a function of school condition and level of transience	181
Table 53: MANOVA output for effect of interaction between school condition and transience level on baseline PYD mean scores	182
Table 54: Mean change scores, standard deviations, and confidence intervals for PYD and CC subscales as a function of transience level x school condition	184
Table 55: One-way RM-ANOVA output for effect of the interaction between school condition and transience level on change in PYD and CC mean scores pre-to-post test	185

## List of Figures

Figure 1: The original five-factor structural model	115
Figure 2: The final three-factor model	116
Figure 3: PYD Model 1	134
Figure 4: PYD Model 2	136
Figure 5: Inter-correlated PYD Model 3	138
Figure 6: The original five-factor measurement model for CC data	140
Figure 7: Four-factor, inter-correlated CC model	141
Figure 8: Model 1: An inter-correlated PYD to CC model	144
Figure 9: Model 2: PYD to CC structural equation model	149
Figure 10: Model 3: CC to PYD model	152
Figure 11: Model 4: A hierarchical PYD to CC structural equation model	155
Figure 12: Model 4A: A hierarchical CC to PYD structural equation model	156
Figure 13: Final autoregressive structural equation model for PYD and CC	161

## Chapter 1: Introduction

He who has health has hope, and he who has hope has everything.

~Arabic Proverb

### 1.1 Child Social Health

Health, and the realisation of a healthy life is a basic human right (United Nations, 1989). Unfortunately, not all children are born into conditions that promote ideal health and wellbeing (Organisation for Economic Cooperation and Development [OECD], 2009; United Nations International Child Emergency Fund [UNICEF], 2014). This is particularly true for children born into poverty. Such children often live in highly stressful conditions characterised by marital discord or conflict (Takeuchi, Williams, & Adair, 1991), harsh and restrictive parenting practices (Jarrett, 1997; Menaghan, 1999; Simons, Johnson, Conger, & Lorenz, 1997), unstable family structures (Lichter & Jensen, 2002), crowded housing (Perry, 2013; Roy, Tubbs, & Burton, 2004), and/or chaotic or disordered living conditions (Vernon-Feagans, Garrett-Peters, Willoughby, & Mills-Koonce, 2011). They may also be forced to move frequently as their families search for employment or affordable housing (Clark, Robinson, Crengle, Herd, Grant, & Denny, 2008; Helu, Robinson, Grant, Herd, & Denny, 2009; Vernon-Feagans et al., 2011).

Over time, exposure to such challenging environments can have a negative effect on children's attitudes, behaviours, and development, particularly in the social dimension (Blane, Berney, Davey-Smith, Gunnell, & Holland, 1999; Davey-Smith, 2003; Marmot & Wilkinson, 2006). Typically, children from impoverished backgrounds also struggle to develop healthy social skills and competencies, establish positive social networks, and forge lasting social relationships with peers and family members (Dodge, Pettit, & Bates, 1994; Elder, Van Nyugen, & Caspi, 1995; Kupermsmidt, Burchinal, & Patterson, 1995; Sameroff, 1986). They are also more likely than their wealthier peers to demonstrate antisocial behaviours (Conger, Ge, Elder, Lorenz, & Simons, 1994; Dishion, French, & Patterson, 1995; Loeber & Stouthamer-Loeber, 1986; McLoyd, Jayaratne, Ceballo, & Borquez, 1994; Sampson & Laub, 1994).

Globally, efforts have been made to combat the negative effects of poverty on social health and development. Many of these efforts have taken the form of community and school-based programmes designed to help children accrue the skills and assets needed to build resilience and promote healthy development. In North America, programmes have typically adopted a positive developmental approach (Bowers, Li, Kiely, Brittan, Lerner, & Lerner, 2010; Catalano, Hawkins, Berglund, Ryan, & Lonczak, 2004; Lerner, 2004, 2007; Lerner, Lerner et al., 2005; Lerner, von Eye, Lerner, Lewin-Bizan, & Bowers, 2010), and focus on providing children with opportunities to develop the *social assets* (i.e., social skills) that they need to realise positive social outcomes – empathy, sympathy, compassion, cultural sensitivity, and respect for others (Bowers et al., 2011; Catalano et al., 2004; Jelicic, Bobek, Phelps, Lerner, & Lerner, 2007).

Such programmes are usually introduced during the intermediate or secondary school years, but some researchers argue that they should be introduced even earlier. Researchers at the Brookings Institute, for example, suggest that interventions targeted at improving social health outcomes for disadvantaged or impoverished children should be offered during the early childhood years (Sawhill & Karpilow, 2014). They argue that impoverished children exposed to targeted, high quality preschool and early primary school intervention programmes will experience success at rates nearly identical to those of higher-income children, and that the benefits last until children reach the end of elementary school, at which point programme participation would have only a modest effect on child behaviour and social skill development (Sawhill & Karpilow, 2014). Research underlining the benefits of early intervention is particularly relevant for New Zealand educators concerned about the high rates of poverty and the poor social health of children in this country.

## **1.2 The Situation in New Zealand**

At present there are 911,300 children aged 0-14, and an additional 246,833 young people aged 15-18, living in New Zealand. In the New Zealand health literature, young people over the age of 15 are typically termed *adults* and health statistics for this group are reported separately from children (Ministry of Health [MOH], 2012, 2013a). For the purposes of this PhD however (and in keeping with international practice for reporting on child health and wellbeing, as mandated by the World Health

Organisation (WHO) and the United Nations (United Nations, 1989)), all young people living in New Zealand aged between 0 and 18 will be termed *children*. Using this criterion, there are approximately 277,952 children (or 24% of the total child population) currently living in poverty in New Zealand (Boston, 2013; Child Poverty Monitor, 2014; Craig, Reddington, Wicken, Oben, & Simpson, 2013; Perry, 2013). This means that New Zealand ranks fifteenth in the international child poverty rankings for developed countries ( $n = 35$ ), a placing corroborated by a 2012 poll conducted by the Innocenti Research Centre which found that New Zealand had a higher rate of child poverty than most Western European and all Scandinavian countries. In fact, only Luxembourg, Portugal, Italy, Greece, and Spain had higher rates. Such results led Boston (2013) to declare that child poverty in New Zealand is an issue that warrants robust study.

The etiology and impact of child poverty in New Zealand are not well understood. Researchers however are contributing to a growing body of literature in this area. In 2014, the Child Poverty Monitor reported that rates of child poverty in New Zealand vary as a function of age, family structure, and ethnicity, and that they are generally higher amongst younger children (< 11 years of age), children from single parent families, and children of Māori and/or Pacific descent. Other studies underline that living in poverty can also be associated with a number of adverse or challenging life experiences. Boston, in a 2013 report on reducing child poverty, argued that impoverished children in New Zealand were more likely to experience chaotic, unpredictable, unsafe, and/or uncomfortable home lives than their wealthier peers, while other researchers report that impoverished children are more likely to move house on a regular basis (Boston, 2013; Child Poverty Action Group, 2014a; Johnson, 2002; Gilbert, 2005), and to experience adverse physical, mental, emotional, and social health outcomes (Adolescent Health Research Group, 2013; Child Poverty Action Group, 2014a; Perry, 2013).

The social health of impoverished children in New Zealand is a particular area of concern because, in New Zealand primary and secondary schools, the rates of antisocial behaviour (bullying, violent behavior, etc.) are increasing (Clark et al., 2013; Denny et al., 2014). In 2012, 9% of the children ( $n = 765$ ) surveyed as part of a national health research study, the Youth 2000 survey series, reported that they were

afraid of being hurt or bothered at school, while 6% ( $n = 510$ ) reported being bullied on a weekly basis (Clark et al., 2013; Denny et al., 2014). In a previous iteration of the same study (Youth '07), 41% of male and 27% of female students had reportedly hit or physically harmed someone else, 26% of males and 12% of females had been in a serious physical fight, and 9% of males and 3% of females had carried a weapon within the previous year (Clark, Robinson, Crengle, Grant, Galbreath, & Sykora, 2009). Of this weapon-wielding population, the study reported that 4% of males and 1% of females had physically attacked someone with the weapon (Clark et al., 2009).

In March 2012, concerns about the high rates of child poverty and poor health and behavioural statistics for this vulnerable population spurred Dr. Russell Wills, the Children's Commissioner in New Zealand, to form an Expert Advisory Group on Solutions for Child Poverty (EAG). The EAG was tasked with compiling local and international evidence about the experiences of children living in poverty, and to make recommendations for pragmatic and effective ways to combat the issue in New Zealand (St. John, 2012). The EAG was comprised of 13 experts from the fields of health, social policy, business, law, and education.

At the same time that the EAG was formed, the national government in New Zealand drafted the first of two policy documents aimed at promoting the health and wellbeing of vulnerable children. The first, *The Green Paper for Vulnerable Children*, called for better protection, child health policy, and service provision for vulnerable children in New Zealand (New Zealand Government, 2012a). It was released to the general public in late 2011, and New Zealanders were asked to engage with it and provide feedback. At the end of this process, *The White Paper for Vulnerable Children* was released (New Zealand Government, 2012b). This new document outlined 12 strategies for improving the health and wellbeing of vulnerable children, including improved childcare provision and child health policy, funding for support services, and child health programme development and evaluation. The strategies formed a comprehensive plan that challenged organisations and individuals to provide vulnerable children with the opportunities, resources, and skills needed to develop into healthy, contributing members of society. At the same time, a companion document, the *Children's Action Plan* (New Zealand Government, 2012b), was also released. This described how the 12 strategies could be realised in practice. The

Green and White Papers subsequently informed the development of the Vulnerable Children's Act and led to the establishment of a Vulnerable Children's Board in parliament. They also provided a mandate for the provision of social health services and programmes that would support vulnerable children to develop healthy social skills, attitudes, and behaviours. Sadly, at the time, understandings about effective social health promotion and proven programme designs, particularly in the school context, did not exist to support the enthusiasm.

In school contexts, social health promotion has traditionally been the purview of the physical and health education teacher, and practice has been guided by the New Zealand Health and Physical Education (HPE) curriculum. The current version of this curriculum was published by the Ministry of Education (MOE) in 2007, and is organised into four strands and seven key areas of learning. Two of the strands – *Relationships with Other People* and *Healthy Communities and Environments*, are directly related to social health education, while the remaining strands and the key learning areas have only an indirect relationship. The stated aims of the HPE curriculum are to support children to develop interpersonal skills and prosocial attitudes, personal values and identity, sensitivity and respect for others, the ability to demonstrate empathy, and the ability to take critical action to promote personal, interpersonal, and societal wellbeing (MOE, 2007). The HPE curriculum is taught in all mainstream primary, intermediate, and secondary schools in New Zealand, but surprisingly little is known about its delivery or implementation. This is particularly true of *low-decile* (i.e., lower socioeconomic area) schools, where accelerating rates of antisocial behaviour do suggest the possibility of delivery issues, and/or the inability of the social health dimension of the curriculum to meet the needs of vulnerable children, and likely, the need for additional or supplementary programming to effect more positive social health outcomes.

In New Zealand, additional or supplementary programming in schools has traditionally been provided by not-for-profit community-based organisations, commonly referred to as *external providers* (Dyson, Gordon, & Cowan, 2011; Petrie, Penney, & Fellows, 2014). In health education, external providers are typically personnel from district health boards, fire, police and ambulance services, local councils, and/or independent health care organisations. To date however, only one

provider has offered a social health promotion programme informed by the positive developmental methodologies that have proven so successful in promoting social skills acquisition in American children, particularly when the provider specialises in that approach (Bowers et al., 2011; Jelicic et al., 2007). In New Zealand, that provider is The Foundation for Youth Development (FYD), a charitable trust based in Auckland.

The social health promotion programme offered by FYD is called *Kiwi Can*, and has been operating in low-decile schools throughout New Zealand since 2003. Like other externally provided health education programmes offered in New Zealand schools however, it has yet to be comprehensively evaluated (Dyson et al., 2011; Petrie et al., 2014). This thesis therefore was designed to evaluate the impact on the social attitudes and behaviours of vulnerable children participating in *Kiwi Can*, as well as the efficacy of competency-based approaches to social health instruction. It was hoped that the results from these evaluations would inform future practice in school-based health education curriculum design and delivery.

Three sequential studies were conducted as part of these endeavours. The first study was an evaluation of the implementation or delivery of the traditional health education curriculum in New Zealand primary and intermediate schools, with a particular focus on the social health dimension. Its' aim was to gather baseline data on implementation that would inform the parameters of the evaluation of the *Kiwi Can* programme. The implementation of the *Kiwi Can* programme was then evaluated, followed by an evaluation of programme outcomes. In this thesis, the first two studies are referred to as *implementation evaluation* studies, while the final study will be referred to as the *outcome evaluation* study.

### **1.3 Significance of Research**

This research, and the three evaluation studies it comprises, is significant for human rights, theoretical, and practical reasons. Firstly, it is driven by a desire to optimise the social health and wellbeing of vulnerable children in New Zealand – an aspiration that aligns with both the principles of the United Nations Convention on the Rights of the Child (United Nations, 1989), and the New Zealand government's desire to support vulnerable children in this country to thrive, achieve, and develop a sense of

belonging (New Zealand Government, 2012a). Secondly, it addresses significant gaps in the health education literature in New Zealand – to date, there is only one published empirical study on the delivery of the health education curriculum in primary or intermediate schools, and no published empirical work on the implementation of externally provided, asset-based, social health promotion programmes in low-decile schools, or the impact that participation in such programmes has on the social attitudes and behaviours of vulnerable children. Thirdly, the results of this research have the potential to affect the way that health education curriculum is designed and delivered in low-decile schools. If the evaluation of Kiwi Can programme outcomes is favourable, then educators and curriculum writers may adopt a positive, asset-based approach to instruction in health education, and/or advocate for the use of external providers to deliver social health programming. This would have a significant impact on pedagogical practice and raise important questions about the role of the health education teacher in New Zealand schools.

#### **1.4 Background to Research**

This thesis grew out of a desire to actively promote improved social health outcomes for vulnerable children in New Zealand, and was inspired by my experiences as a physical and health education teacher in low-decile schools in North America, Europe, the United Kingdom, Asia, and Oceania. In each of these countries I found that vulnerable children faced similar social challenges around peer group assimilation and identity, and were often ill equipped to cope with aggressive and antisocial behavior. Often such children failed to thrive, were emotionally isolated and depressed, and in some cases suicidal. As a health education teacher I felt that it was my responsibility to help these children develop the skills needed to deal with the instability and social complexities of life, and particularly life in low-decile schools. I advocated for the delivery of social health education that was competency-based, believing that the most effective way to enhance children's social health was to equip them with practical skills that could be used in real life situations. Unfortunately physical health was invariably the primary focus in all the countries and schools where I taught. Priority was given to regular fitness training and maintaining a healthy body weight and a healthy lifestyle – often to the exclusion of other aspects of wellbeing, particularly those with a social dimension. As my career progressed I

began to suspect that the health curriculum in schools was driven and shaped by political concerns around the financial costs of obesity and ill health, and was excited to encounter social health programmes in New Zealand, albeit externally provided. I became interested in the potential of such programmes to positively influence the social attitudes and behaviours of vulnerable children, and resolved, despite my secondary teaching background, to focus my research endeavours on younger children, believing that earlier intervention would likely yield the most significant and sustained benefits. I then embarked on the three-year research project described in this thesis.

## **1.5 Assumptions**

This research is underpinned by a number of assumptions about the nature and plasticity of children, the course of human development, the importance of health promotion activities, the holistic nature of health, the purpose of school-based health education, and the role that health educators can play in improving child health and well-being. Specifically, it is based on assumptions inherent in developmental systems theory: (i) that every child has the potential to develop positive social attitudes and behaviours; (ii) that positive social attitudes and behaviours are associated with positive social health outcomes; and (iii) that children can be supported to develop the skills they need to enjoy improved social health and development (Benson & Scales, 2009). It is also based on the assumptions that a primary goal of health education, whether delivered by classroom teachers or external providers, is to equip students with the skills required to realise positive social health gains, and that social attitudes and behaviours will influence children's participation in social health programmes.

## **1.6 Structure**

This thesis is presented in seven chapters. Chapter 1 overviews the research, discusses its significance, provides information about the researcher's background, and identifies key underlying assumptions.

Chapter 2 reviews the literature, situates the research in a historical context, and elaborates further on constructs of interest to this study – child poverty and

vulnerability, transience, social health promotion, and child development. Chapter 2 also includes a discussion of positive youth development, the main theory underpinning the research, and presents common tools used to measure change in social attitudes and behaviours at both the individual and group (classroom) level.

Chapter 3 discusses the general design and theoretical methodologies used by the researcher, and provides a rationale for the use of the various quantitative survey techniques employed.

Chapter 4 discusses the design, methodology, and results of the evaluative study of traditional health education curriculum delivery in New Zealand primary and intermediate schools, with a particular focus on teacher beliefs about health education, curriculum implementation, and perceived barriers.

Chapter 5 provides a detailed overview of the Kiwi Can programme, and then discusses the design, methodology, and results of the Kiwi Can implementation evaluation study and its implications for outcome evaluation design.

Chapter 6 presents a discussion of the design, methodology, and results of the Kiwi Can outcome evaluation study, and a discussion of the design and testing of structural equation models created to capture the relationships between factors at the individual and classroom level, subsequently used to inform the analysis of mean differences in social attitudes and behaviours between Kiwi Can programme and non-programme (control) participants both within and across time. Chapter 6 also presents and discusses the results of a small exploratory study on the effect of transience on self-reported prosocial attitudes and behaviours amongst student participants.

Finally, Chapter 7 summarises the results from the outcome evaluation study and discusses the implications for social health programme design and delivery in low-decile schools. It concludes with a discussion of the inherent limitations of this research and makes suggestions for future research.

Copies of all of the measurement tools, Kiwi Can curriculum documents, ethics paperwork, and other supporting documentation for this thesis are included in the appendices. A Memorandum of Understanding dated November 2011 and signed by

the Kiwi Can programme team and attesting to their support for this research is also appended (see Appendix A).

## **Chapter 2: Literature Review**

### **2.1 Introduction**

This thesis examines positive youth development programmes in low-decile primary and intermediate schools in New Zealand, programmes that target those New Zealand children most at risk of negative life consequences, and therefore the children most in need of effective health education interventions. Literature that is contextually relevant for this thesis, and for the research that follows, is embedded in this chapter. Firstly, the nature of child poverty in New Zealand is defined, and the risks faced by children from impoverished homes are examined. Secondly, the need for health promotion curricula and programmes within schools is established, and the historical shift within health education from disease prevention to health promotion is examined, as are models of health promotion, and the social factors that contribute to health quality. Important psychological theories are also considered to elicit thinking and strategies around personal responsibility for effective health outcomes, and how programme design and delivery has embraced these ideas. Thirdly, an overview of the New Zealand schooling system is provided, and the impact of childhood vulnerability on schools is discussed, with a particular focus on transience. The social climate of schools and its measurement is then reviewed, recognising the significant relationship between health outcomes and children's social experiences. The chapter concludes with a discussion of both the health education curriculum in New Zealand primary and intermediate schools and the practice of outsourcing social health education programmes. It is followed by a presentation of the research questions that guided the empirical work described in this PhD thesis.

### **2.2 Child Poverty in New Zealand**

In 2012, the EAG defined children living in poverty as those deprived of the material resources and income needed to develop and thrive, children who were therefore unable to enjoy their rights, achieve their potential, or participate equitably as members of society.

At present, there are approximately 1.2 million children living in New Zealand, and at the time of the last census (2013), Statistics New Zealand reported that 72.4% of these

children were Pakeha/European, 23.7% were Māori, 12.1% were Pasifika, 9.7% were Asian, and 1.2% identified as *Other* (Lang, 2002; Statistics New Zealand, 2014b). In New Zealand 24% of the child population currently live in poverty (Craig et al., 2013; EAG, 2014), and of these children, 60% live in persistent poverty (seven or more years lived in chronically impoverished conditions), and 10% live in extreme poverty (Perry, 2014).

In New Zealand, poverty is typically defined at the household level and based on either a relative or fixed-line measure of disposable income. In both measures, a poverty benchmark is established and household incomes are compared to the benchmark. The benchmark for poverty in New Zealand – when a household is deemed to be impoverished is when household income is less than 60% of the median household income, net-of-housing costs in the current year (termed *relative poverty*), or at a fixed point in time (termed *fixed-line poverty*). Relative measures of poverty typically use median income rates from 2008, whereas fixed-line measures use rates from either 1998 or 2007 (Children’s Commission, 2012; Perry, 2011). For this thesis, measures of relative poverty will be used, in line with current reporting practices in New Zealand (EAG, 2014).

In New Zealand, relative poverty is measured using The New Zealand Deprivation Index (NZDep). The NZDep is a measure of relative regional socioeconomic deprivation, which rates small geographical areas, or *meshblocks*, against one another within the national context (Atkinson, Salmond, & Crampton, 2014; Salmond, Crampton, & Atkinson, 2007). Each meshblock is assigned a deprivation score calculated using national census data related to income, home ownership, employment, qualifications, family structure, housing, access to transport, and communication. Scores range from 1 to 10. A score of ‘1’ is assigned to areas with the lowest levels of deprivation, and scores of ‘10’ are assigned to areas with the highest levels of deprivation (Atkinson et al., 2014; Salmond et al., 2007). Ten percent of the meshblocks are assigned to each deprivation level, and deprivation scores are re-calculated after each census.

The NZDep 2013 established no clear geographical pattern of deprivation in New Zealand. Areas of high deprivation were noted as being scattered throughout the country, though they were more prevalent in the North Island (Atkinson et al., 2014).

NZDep 2013 also noted that the demographic composition of high deprivation areas varied (Atkinson et al., 2014; Dale, O'Brien, & St. John, 2011), though Māori and Pasifika families were typically overrepresented compared to Pakeha, Asian, and Other families. In 2013, a report by the Adolescent Health Research Group quantified the disparity, reporting that 46.2% of Māori children and 73.4% of Pasifika children lived in areas of high deprivation (NZDep 2006 levels 8-10), compared to only 15.1% of Pakeha, 27.5% of Asian, and 28.0% of Other children. In 2014 a report by The Child Poverty Monitor corroborated some of these figures, stating that between 2010 and 2013, 34% of Māori and 28% of Pasifika children lived in poor households, compared to only 16% of European children. Though none of these reports explored the reasons for the disparities identified, similar studies in the United Kingdom (Barnard & Turner, 2011), North America (Government of Canada, 2007; Gradin, 2008), and South-East Asia (Baulch, Nguyen, Phuong, & Pham, 2010) offer explanations that may be relevant to the New Zealand experience. These studies argue that higher child poverty rates in ethnic minority populations are the result of higher rates of unemployment, lower levels of educational attainment, and disparities in salaries and wages. All the studies also noted however that the etiology and causes of poverty vary dramatically both within and between ethnic minority groups.

While the causes of poverty in ethnic minority (and majority) populations are complex and varied (Garrett, Ng'andu, & Ferron, 1994; Korbin, 1992), researchers argue that the lived experiences of children living in poverty have markedly similar characteristics – most will experience instability, unpredictability, and/or chaos (Lichter & Jensen, 2002; Vernon-Feagans et al., 2011), many will live in cold, damp houses that are overcrowded and poorly insulated, and be persistently hungry (Boston, 2013), and many will move homes with their families at regular intervals (Boston, 2013; Clark et al., 2008; Helu et al., 2009; Vernon-Feagans et al., 2011).

### **2.3 Transience**

The term *transience* first emerged in the United States in the 1960's and 1970's, and was used to describe the movement of military personnel, professional groups, and seasonal workers who relocated frequently (Bruno & Isken, 1996; Edwards, 1997; Fields, 1997). Over time the concept of transience expanded, and researchers adopted the term to refer to the frequent movement of any group, including children, from one

location to another. Studies of the incidence, causes, and impact of transience in childhood have been undertaken in the United States (Ackerman, Kogos, Youngstrom, Schoff, & Izard, 1999; Alexander, Entwisle, & Dauber, 1996; Hanushek, Kain, & Rivkin, 2004; Kerbow, 1996; Schacter, 2001a, b), the United Kingdom (Boheim & Taylor, 2002; Brown, Benzeval, Gayle, Macintyre, O'Reilly, & Leyland, 2012; Dobson & Henthorne, 1999; Dobson, Henthorne, & Lynas, 2000), Australia (Fields, 1997; Hill, Navin, & Lynch, 2009; Sorin & Ilostes, 2003), and Canada (Koschmider, 1997). In these studies there is a general consensus that transience is primarily influenced by changes in family circumstances (e.g., separation or divorce), changes in employment circumstances (e.g., redundancy or transfer), and housing issues (e.g., rent increases or the sale of rental accommodation). Researchers have also established that transience is more likely amongst families who reside in areas where accommodation options are less desirable or shorter term (Dobson, Henthorne, & Lynas, 2000), and when they are from low-socioeconomic or ethnic minority backgrounds (Hanushek, Kain, & Rivkin, 2004; Rumberger, Larson, Ream, & Palardy, 1999; Waldegrave, 2000; Waldegrave & Stuart, 1998). Rates of transience differ significantly between countries. Scandinavian countries typically record the lowest rates of transience, while countries in Eastern Europe and New Zealand are amongst those with the highest rates (Jelleyman & Spencer, 2008).

### **2.3.1 Transience in New Zealand.**

In a recent international comparison of population mobility in 35 developed countries, rates of transience in New Zealand were 2.1% higher than those in the United States, 2.6% higher than those in Australia, and 5.0% higher than those in the Western European and Scandinavian countries surveyed (Innocenti Research Group, 2012). Transience in New Zealand is strongly driven by housing availability and the expense of purchasing residential property (Collins, 2000; Johnson, 2002; Kelsey, 1998; Waldegrave, 2000). In the early 1990s, changes were made to the policies governing rental accommodation owned and managed by Housing New Zealand, a crown agency. These changes were the catalyst for significant increases in private sector rents, and rates have risen steadily ever since (Collins, 2000; Perry, 2010) forcing many New Zealand families to adopt a transient lifestyle in an effort to find and remain in affordable housing and control debt (Johnson, 2002). The search for paid

employment, family breakdown and dysfunction, and the desire to evade social service personnel, have also been reported as significant causes of transience in New Zealand (Neighbour, 2002). Youth '12, a national survey of adolescent health, underlined that there is now a generation of New Zealand children who are routinely uprooted from their homes. In 2012, 7% of study participants reported moving home two or more times in the past 12 months, and movement was reported as significantly more common (10%) in communities and schools in high deprivation areas (NZDep 8-10), than in communities and schools in middle (NZDep 4-7) or low (NZDep 1-3) deprivation areas (Clark et al., 2013).

### **2.3.2 The education system in New Zealand.**

The schooling system in New Zealand has four sequential levels – early childhood education, primary schooling, secondary schooling, and tertiary education (MOE, 2014a). Education is free for all New Zealand citizens and permanent residents between the ages of five and 19, and is compulsory for all students between the ages of six and 16 (MOE, 2014a). The state school system offers classes from the new entrant level (Year 0) to Year 13, and most students will complete all 13 years of schooling in either single-sex or co-educational schools. At present, 85% of students in New Zealand attend state schools, 11% attend integrated schools (effectively state schools, but with a specific religious or pedagogical approach that guides practice), while 4% attend private schools, boarding schools, schools that cater for special educational needs, or are home-schooled (MOE, 2014a). Primary schools that offer classes for Years 1-6 are called *contributing primary schools*, while those offering classes for Years 1-8 are called *full primary schools*. Schools that offer classes for Year 7 and 8 only, are called *intermediate schools*, while those that offer classes for Years 1-13 are called *composite schools* (Education Counts, 2014). As of July 2014, there were 1,078 full primary schools, 766 contributing schools, 117 intermediate schools, and 162 composite schools in New Zealand (Education Counts, 2014).

Each school has a designated home zone, and all children who reside within that zone are guaranteed enrolment (Te Kete Ipurangi [TKI], 2014). Home zones vary in size and density (TKI, 2014), and schools can be affiliated with more than one zone. Each school is also given a decile rating that reflects the socioeconomic status of its zone (NZDep meshblock scores are used in the calculation), and there are 10 decile levels.

Each level comprises approximately 10% of the schools in New Zealand. A decile rating of '1' is given to the schools that have the highest proportion of children from low-socioeconomic communities, while a decile rating of '10' is given to schools that have the lowest proportion of children from low-socioeconomic communities (MOE, 2014b). Decile ratings are re-calculated after every census to ensure that they accurately reflect the school's community, and were most recently updated in November 2014 (MOE, 2014c).

### **2.3.3 Transience in New Zealand schools.**

In New Zealand, transience is defined at the individual and at the school level. Individuals are deemed to be transient if they have attended two or more schools in a year (Education Review Office, 2007; New Zealand Center for Educational Research, 2005), while transience in schools is determined by the number of children who join or leave at non-normal entry or exit points (Neighbour, 2000), generally using a formula developed by Johnson in 2002 for a report on mobility in South Auckland primary schools. In this report school level transience was calculated according to the following formula:  $[\text{enrolments during the year} - \text{increase in school roll}] / \text{average school roll for the year}$  (Johnson, 2002, p. 3).

Transience in New Zealand schools is a significant issue. Children have been known to change schools as often as ten times over the course of a seven year primary school career (ERO, 2007), and in 2013, the MOE reported that there were 3,744 individual transient school children in the previous year. Of those children, 87% transferred schools twice, 11% transferred three times, and 2% transferred to four or more schools. Johnson's 2002 study involving 59 South Auckland primary schools reported an average school level transience rate of 29%, and that 32% of the children living in Decile 1 areas (determined using the NZDep index) and 17% of children living in Decile 3+ areas had changed schools. She also reported that families with the lowest 10% of incomes were twice as likely to move as children from higher income households. Other studies and reports corroborate these results and further underline that transience is an ongoing issue for all New Zealand schools, irrespective of location and demographics, but that it is a particular issue for low-decile schools.

In 1999, Wylie's national survey of New Zealand schools reported an average rate of transience of 26%, and that the rate varied as a function of school decile (low-decile schools in the survey averaged 42%, while in high-decile schools the average rate was only 7%). Neighbour's 2000 survey of every primary and intermediate school in New Zealand ( $n = 795$ ) reported an average transience rate of 30%, but a rate of 43% in Decile 1 schools, and 26% in Decile 10 schools. Gilbert's 2005 study of 19 schools in rural North Island areas reported a mean school-level transience rate of 35%, and rates that varied from 9% to 45% (Gilbert, 2005; Gilbert & Bull, 2007). More recent studies by McLeod (2014) in Hamilton and Porirua, also reported significantly higher rates of transience in low-decile schools, including one school ( $n = 167$ ) where the rate of transience in 2013 was 55.1%, and where 50 new students joined the roll and 33 departed in the first month of the following school year. Transience in low-decile schools has also been explored by the Child Poverty Action Group (CPAG) and by the MOE. In a 2014 media release, CPAG reported that Decile 1 schools in South Auckland had experienced a median roll turnover of 30% during 2013 (CPAG, 2014b), while in the same year the MOE reported that the average rate of transience in Decile 1 primary schools nationally was 52.8% (compared to a rate of 29.4% in Decile 10 schools). When rates were also compared between students as a function of ethnicity, Māori children were found more likely to be transient than their non-Māori peers (MOE, 2013).

### ***2.3.3.1 Impact of transience on classroom climate.***

The movement of students in and out of a classroom can have a significant, negative impact on classroom climate (Sorin & Iloste, 2003). Researchers have established that antisocial classroom behaviour by some children can erode the sense of safety or belonging that other children feel, and that such feelings are exacerbated in highly transient classrooms (Bradshaw, Sawyer, & O'Brennan, 2009; Wilson, 2004). In a study of highly transient children in American schools, McLeod (2014) found that children who relocate frequently from one school to another are often aggressive and hostile toward their peers, and lack the skills, particularly the language skills, to cope effectively in new social situations or to resolve conflicts peacefully. Wilson (2004) reported that in extreme cases, the behavior of such children was threatening, and even dangerous. The regular movement of children in and out of classrooms also

disrupts established routines (Bruno & Isken, 1996; Fisher, Matthews, Stafford, Nakagawa, & Durante, 2002; Gilbert, 2005; Lee, 2000; Neighbour, 2003), and can force teachers to re-configure classes in ways that destabilise instructional continuity, and force children to redefine their friendship groups in ways that upset or disturb the existing classroom dynamic (Sorin & Iloste, 2003). At the school level, transience can be equally destabilising and disruptive. In studies on transience and school climate in the United Kingdom, Europe, and North America, high rates of student mobility were associated with an increased risk of school violence (Birnbaum, Lytle, Hannan, Murray, Perry, & Forester, 2003; Khoury-Kassabri, Benbenishty, Astor, & Ziera, 2004; Stewart, 2003) and a diminished school climate (Bevans, Bradshaw, Miech, & Leaf, 2007; Griffith, 2000; Neighbour, 2003; Payne, Gottfredson, & Gottfredson, 2003).

### ***2.3.3.2 Measuring classroom and school climate.***

Historically, self-report surveys have been used to measure children's perceptions of the psychosocial dimensions of their classroom environment (Fraser, 1980, 1981, 1982; Moos, 1974), and, in a few cases, those of their teachers as well (Fraser, 1981; Rentoul & Fraser, 1979). Such measurement instruments typically varied in length and item-response format, and most were designed for use with secondary school students (e.g. the Learning Environment Inventory [Fraser, 1981], the Classroom Environment Scale [Fraser & Fisher, 1986; Moos, 1974], and the Individualized Classroom Environment Scale [Rentoul & Fraser, 1979]). Interest in measuring the climate of primary school classrooms has however grown significantly in the past two decades, and in Australia and America researchers have used the My Class Inventory survey to do so. The My Class Inventory survey is a simplified version of the Learning Environment Inventory, and was originally designed for use with children aged eight to 12 (Fraser, 1982; Fraser & Fisher, 1986). It has five scales – Cohesiveness, Friction, Difficulty, Satisfaction, and Competitiveness, and uses a simple 'Yes-No' response format (Fraser & Fisher, 1986). Critics have however regularly questioned the reliability of three of the scales (Cronbach  $\alpha = \leq .60$ ) (Fraser, Anderson, & Walberg, 1982), and also bemoaned the tool's narrow scope and inability to capture information on the social attitudes and behaviours of children, or

their social interactions. For these reasons the My Class Inventory survey was not selected for use in this thesis.

Instead, a tool was selected that was capable of measuring the social climate of school classrooms following participation in a social health promotion programme. That tool is an adaptation of the School Climate Survey created at the Developmental Studies Centre in the United States by researchers responsible for evaluating the impact of participation in school-based child development projects (Solomon, Battistich, Kim, & Watson, 1997). The research team there defined school climate as, “the quality of the social relationships among school members” (Roberts, Hom, & Battistich, 1995, p. 2), and designed a measurement tool to assess it in public schools across six districts. The original tool had 11 factors and over 100 items, and demonstrated adequate factor structure ( $CFI = \geq .83$ ) and reliability (Cronbach’s  $\alpha = .70 - .92$ ) (Solomon, Battistich, Watson, Schaps, & Lewis, 2000), but practical limitations associated with its use, particularly in relation to funding and time constraints. An abbreviated version of the survey was then created for use in measuring the climate of schools engaged in a positive development and character building programme (EPIC Pathways to Character) in Buffalo (Ding, Liu, & Berkowitz, 2011), and it is this tool that was selected for use in this thesis. The abbreviated version had seven factors – Positive Behaviour ( $n = 10$  items, Cronbach’s  $\alpha = .82$ ), Negative Behaviour ( $n = 5$  items, Cronbach’s  $\alpha = .78$ ), Classroom and School Supportiveness ( $n = 5$  items, Cronbach’s  $\alpha = .78$ ), Autonomy and Influence ( $n = 3$  items, Cronbach’s  $\alpha = .62$ ), Safety at School ( $n = 2$  items, Cronbach’s  $\alpha = .64$ ), Enjoyment of Class/School Liking ( $n = 4$  items, Cronbach’s  $\alpha = .74$ ), and School Norms and Rules ( $n = 5$  items, Cronbach’s  $\alpha = .70$ ). The factor structure was tested using exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and structural equation modeling (SEM) techniques. The results of the SEM indicated that the null model with the seven factors and 34-items had adequate fit –  $CFI = .96$ ,  $RMSEA = .017$ . It also demonstrated measurement and scalar invariance ( $\Delta CFI = .02$ ). Formal permission was sought and granted for the use of the abbreviated tool in this research.

## 2.4 Poverty, Transience, and Social Health

The relationship between childhood poverty, transience, and children's social health is effectively illuminated by the literature, albeit circuitously. The impact on non-transient children who attend transient schools however, is largely overlooked.

### 2.4.1 Poverty and social health.

Poverty is correlated, if not causally linked, with a raft of poor physical, mental, emotional, and social health outcomes for children, and when children live in persistent poverty, the effects are particularly pronounced (Aber, Bennett, Conley, & Li, 1997; Korenman, Miller, & Sjaastad, 1995). In the international literature, the relationship between childhood poverty and poor health outcomes is well established. Davey-Smith (2003) reports that childhood poverty is an independent predictor of injuries and deaths related to heart and respiratory diseases, diabetes, and some cancers. Childhood poverty is specifically associated with physiological damage and an increased incidence of premature death (Craig, Jackson, & Han, 2007; Mare, 1982; Marmot & Wilkinson, 2006). Impoverished children are reported at greater risk of poor mental health outcomes (Brown, Bhrolchain, & Harris, 1975; Liem & Liem, 1978; Warheit, Holzer, Bell, & Arey, 1976), impaired cognitive development (Aber et al., 1997; Duncan, Brooks-Gunn, & Klebanov, 1994; Schoon, Jones, Cheng, & Maughan, 2012), and failure to form healthy emotional attachments with parents and caregivers (Duncan et al., 1994). They are also more likely to have problematic peer relationships (Dodge, Pettit, & Bates, 1994; Kupermsmidt, Burchinal, & Patterson, 1995), experience social isolation (Lockwood, 2002), present with low levels of social competence (Elder et al., 1995), demonstrate antisocial behavior (Conger et al., 1994; Dishion et al., 1995; Loeber & Stouthamer-Loeber, 1986; McLoyd et al., 1994; Sampson & Laub, 1994), belong to *negative networks* that threaten their safety and ability to develop healthy social connections (Belle, 1983), experience unhealthy, dangerous, and illicit behaviors, and to be the victims of violence themselves (Merry, 1981). For children living in persistent poverty, the negative social health effects can be particularly pronounced (Aber et al., 1997). In a study on the effects of poverty on the health of children in the United States, Korenman, Miller, and Sjaastad (1995) found that children in the *long-term* poor category reported larger deficits in socioemotional development than the *non-poor* children in the sample.

In New Zealand, researchers have presented similar findings on the impact of childhood poverty on children's physical health and wellbeing, but the impact on social health has been largely overlooked. In fact, Easton and Ballantyne (2002) report that New Zealand children living in poverty are three times more likely to be sick than their advantaged peers. Perry (2013) reports that impoverished children are 1.4 times more likely to die during childhood, 1.5 times more likely to be hospitalised, and three times more likely to die from injury. Other researchers have noted that impoverished children in New Zealand are also at greater risk of contracting infectious diseases, living with disabilities, experiencing heart disease, being alcohol and drug dependent, and experiencing poor oral health in adulthood than their higher-income peers (Boston, 2013; CPAG, 2014a).

#### **2.4.2 Transience and social health.**

The link between transience (an indicator of poverty) and poor social health outcomes is well established in the literature. Empirical studies conducted in the United Kingdom and United States report that highly transient children struggle to form meaningful social relationships with others (Carlo, Fabes, Laible, & Kupanoff, 1999; Eisenberg, 1996; Kariuji & Nash, 1999; Rumberger, Larson, Ream, & Palardy 1999; Sanderson, 2003; Simmons, Carlton-Ford, & Blyth, 1987; Strand, 2000; Vail, 1996), face greater risk of failing to form positive social attachments, and higher rates of peer rejection (Hartman, 2002; Hodges, Boivin, Vitaro, & Bukowski, 1999; Rigby 2000, 2001). Highly transient children are also more likely to experience bullying and social alienation (Lloyd, Stead, Jordan, & Norris, 1998; Sorin & Iloste, 2003), and to act in antisocial ways that adversely affect their ability to forge healthy social relationships and establish membership in positive social networks (DeWit, Offord, & Braun, 1998; Simpson & Fowler, 1994).

These findings are echoed by researchers in New Zealand. Utter and Denny (2010) report that transience is a strong, independent predictor of low social cohesion and networking, and that transient children in New Zealand are more likely to report poorer social relationships with neighbours and lower levels of safety. Other researchers report that transient children are more likely to experience difficulties making friends and integrating socially than their less transient peers (Gilbert, 2005; Lee, 2000; Neighbour, 2000, 2002).

### 2.4.3 Gender, adolescence, and social health.

Fisher et al. (2002) argue that the antisocial behavior of transient children reflects a lack of attachment and bonding with peers and adults, while Sorin and Iloste (2003) contend that transient children engage in antisocial behaviour because they feel marginalised and socially isolated, and so befriend other marginalised children who may also be engaged in antisocial behaviour. Irrespective of cause, the incidence of antisocial behavior is significantly higher if the transient child is male or adolescent (Archer, 2009; Jolliffe, 2013; Toldos, 2005; Tsorbatzoudis, Travlos, & Rodafinos, 2013; von Collani & Werner, 2005).

Experience and socialisation are generally proffered to explain gender differences in self-reported antisocial behavior amongst transient children. Researchers argue that boys are more consistently exposed to, and involved in, physically aggressive, and antisocial behaviours (Archer, 2009; Toldos, 2005; Tsorbatzoudis, Travlos, & Rodafinos, 2013; von Collani & Werner, 2005), while girls are more likely to learn via their interactions that empathic, caring, and compassionate behaviours are anticipated and socially acceptable. Henshaw, Kelly, and Gratton (1992) argue that after continued exposure to the same gendered message, children as young as eight or nine will develop gendered understandings, classify prosocial behaviours as *feminine* or *masculine*, and behave accordingly.

In adolescence, children experience a rash of biological, physical, and emotional changes that can affect their ability to engage positively with peers (Garbarino, 1985; Lerner & Castellino, 2002). Children often feel disengaged from school and experience negative relationships with their peers (Akos, 2002; Rice, Frederickson, & Seymour, 2011). A large, longitudinal study in Australia on the social health effects of the middle school transitional period, reported that 31% of the Grade 8 children surveyed ( $n = 1500$ ) experienced difficulties in their final year, and apprehension about the upcoming transition to high school, and that the majority of those children reported significantly lower mean scores for connectedness to school, feelings of safety at school, and prosocial tendencies at the end of the Grade 8 year than at the beginning. They also reported higher mean scores for victimisation and conduct problems (Waters, Lester, Wenden, & Cross, 2012).

#### **2.4.4 Non-transient children in transient schools.**

The impact of poverty and transience on the social health of transient children is widely explored in the literature, but scant attention has been paid to children who attend highly transient schools, but are not themselves transient. These children are typically referred to as *immobile* or *non-mobile* children (Alexander, Entwistle, & Dauber, 1996; Gibbons & Telhaj, 2011). These non-mobile children are unique because they experience the same insecure environments as transient children, but are viewed through a lens of stability. Further, they are often overlooked in the literature because they are not faced with the changes and challenges that confront their transient peers (Gibbons & Telhaj, 2011; Hanushek, Kain, & Rivkin, 2004), or they are regarded as a control group whose lived experiences are not relevant to research objectives (Wood, Halfon, Scarlata, Newacheck, & Nissim, 1993). In many cases, however, non-mobile children face similar disruptions to their social networks, and similar challenges adjusting to the changes wrought by transience, as well as the need to include and induct the new arrival (Felner & Adam, 1988; Ladd & Price, 1987). In a rare study in American secondary schools that focused on the impact of transience on non-mobile children, researchers discovered a correlation between dropout rates and highly transient school environments (South, Haynie, & Bose, 2007). In this study the researchers did not explore the impact of highly transient schooling environments on the social attitudes and behaviours of non-mobile or *non-transient* students. This omission is typical of the literature.

The current thesis attempts to address this gap by exploring differences in the self-reported incidence of prosocial attitudes and behaviours amongst non-mobile students participants of a social health promotion programme operating in low-decile primary and intermediate schools. The programme, Kiwi Can, is based on a positive, health promotion approach.

#### **2.5 The History of Social Health Programming and Education**

The field of health promotion is relatively new, and was preceded by centuries of health education and health programming informed by a biomedical, disease prevention orientation to child health and wellbeing. A brief history of the two

theoretical approaches is provided below, followed by a discussion of positive developmental theories as they pertain to social health promotion in children.

### **2.5.1 Disease prevention approach.**

In the disease prevention model, health is an objective state characterised solely by an absence of biological disease markers (Engels, 1977). Individuals are healthy if they don't exhibit any signs or symptoms of pathology. It is a biomedical model, strongly influenced by germ theory and the work of Pasteur and Koch in the nineteenth century (Engels, 1977). Their groundbreaking experiments identified agents responsible for causing illness and death, and gave rise to the doctrine of specific aetiology – “for every disease there is a single and observable cause that can be isolated” (Ottewill & Wall, 2004, p. 13). The biomedical model of health is reductionist in nature and assumes a mind-body dualism that is informed by a scientific research paradigm (Engels, 1977). It argues that the role of health care researchers and practitioners is to recognise and detect the signs and symptoms of disease and administer appropriate treatment. Practitioners are also expected to educate the public about the potential causes of ill health, and work with professional organisations to prevent disease, through the provision of targeted behavioural interventions (Bristol, 2012).

In the biomedical model, health is viewed as a unidimensional, physical trait that is only assumed to exist when symptoms of biological disease are absent. Historically, it has been a view, much criticised by scientists, researchers, educators, and medical specialists who felt that ‘health’ was a more holistic construct, as championed by the WHO. The WHO definition of health was formally published in 1948. It defined health as a:

state of complete physical, mental, and social well-being, and not merely the absence of disease or infirmity. (Nutbeam, 1998, p.1)

The WHO definition marked the first time that health was defined on the international stage in broad, multidimensional terms (Lindstrom & Eriksson, 2006). Although it has been criticised for being too utopian, too broad, and too difficult to operationalise and measure (Bowling, 2005), it has nonetheless stimulated rigorous discussion and

debate about the nature of health and wellbeing, and created the space for new, and alternative theories and models to emerge.

One of the most popular models to emerge was a biopsychosocial, integrated health model developed by American psychiatrist, George L. Engels in 1977. Concerned with the inherent limitations of the negative, reductionist, biomedical model, Engels proposed an alternative approach for studying health and disease, in which the biological (physical), psychological (mental), and social dimensions of health were presented as distinct entities that were inter-related (Engels, 1977). Engels posited that the three dimensions intersected and that change in one area could produce change in the other two. His ideas built upon the WHO's definition of health, gave rise to new understandings about the importance of study in the mental and social health areas, and spurred researchers, educators and clinicians to establish clear definitions for these constructs. One of the most commonly used definitions for social health was advanced by Russell in 1973, who argued that social health is:

that dimension of an individual's wellbeing that concerns how he gets along with other people, how other people react to him, and how he interacts with social institutions and social mores. (Russell, 1973, p. 75)

It is a broad definition that encompasses elements of personality and social skills (Bowling, 2005; McDowell & Newell, 1996), makes reference to social institutions, and alludes to the impact of cultural or societal norms on social health and wellbeing.

Given the broad nature of definitions such as Russell's, it is unsurprising that social health has been operationalised and measured in so many different ways. At the individual level, social health is often viewed as a function of social and community adjustment, and social support and networking (McDowell & Newell, 1996). Social adjustment refers to the social skills an individual has to function adequately in personal relationships, while social support refers to the emotional and physical support they receive from others. Social networks are considered to be a subset of social support, and refer to the social relationships that an individual has and can rely on. At the ecological level, social health is typically defined and measured as a function of social cohesion (McDowell & Newell, 1996). Social cohesion refers to connectedness and solidarity between groups of people (Kawachi, Kennedy, & Glass,

1999). It is typically measured using tools that question individuals about their feelings of commitment, trust, and belonging (Bowling, 2005). In this thesis, social cohesion (intrapersonal relationships) and social adjustment (individual social skills) of children are used as proxy measures of individual social health and wellbeing for vulnerable children involved in social health programmes. Most of these programmes now adopt a health promotion approach.

### **2.5.2 Health promotion.**

The field of health promotion arose in the 1970s. At the same time that Engels (1977) published his biopsychosocial model, Mr. Marc LaLonde, then Minister of National Health and Welfare in the Canadian federal government published a report, *A New Perspective on the Health of Canadians*, which outlined a fresh vision for health and the health field in that country. Lalonde argued that health was influenced by many factors beyond the individual (human biology, environmental factors, lifestyle choices, etc.), and that it could be improved through the introduction of healthier public policy and practice (Lalonde, 1974). The report heralded the beginning of a new movement in public health internationally (now referred to as the *post-modern, health promotion movement*), and a new focus on the development and mobilisation of health resources that would enable individuals to live healthy and productive lives (Lindstrom & Eriksson, 2006). The principles of the new movement were formally drafted at the first international health promotion conference (held in Ottawa in 1986), and subsequently enshrined in The Ottawa Charter for Health Promotion (1986). In the Charter, health promotion was defined as:

the process of enabling individuals to increase control over and to improve their health, in order to reach a state of complete physical, mental, and social wellbeing, and to lead an active and productive life – that is, a good quality of life. (Eriksson & Lindstrom, 2007, p. 938)

For the first time health was viewed as a resource (rather than as an objective for living), and the notions of caring, holism, and ecology were given voice (WHO, 2009). At the individual level, the charter stated that the prerequisites for health and health promotion were peace, shelter, education, food, income, a stable ecosystem, sustainable resources, social justice, and equity. At the social, environmental, and

governmental level, priority was given to health promoting actions such as building healthy public policy, creating supportive environments, strengthening community action, developing personal skills, and re-orientating the health services sector from a disease prevention to a health promotion focus (WHO, 1986, 2009). It was a paradigm shift, away from disease prevention, and toward health promotion.

In the past twenty years, the field of health promotion has grown rapidly (WHO, 2009). Additional charters have been drafted and countries around the world have committed resources to promoting the health and wellbeing of their citizens through targeted health promotion interventions (WHO, 2009). Many of these programmes have adopted a developmental approach underpinned by bioecological models of health and development. The most widely cited of these models is Bronfenbrenner's bioecological model of human development (Bronfenbrenner, 1979). In Bronfenbrenner's model, individual development is the result of bi-directional interactions between the child and factors present in the social context or ecology surrounding the individual. These factors are termed *structures* and are grouped into five levels. Each structure represents a different level of influence on development. The strongest and most proximal influences are those present in the *microsystem*, or the environment immediately surrounding the individual (school, peers, and family). Bronfenbrenner (1979) argues that factors in the microsystem become richer and more complex as children grow, and that they interact in a bi-directional manner with features at all levels of the ecological model. All interactions between these factors in the microsystem are represented in the second level of model – the *mesosystem*, which, in turn, is surrounded by the *exosystem*. The exosystem includes those features of the environment that indirectly influence development through interactions with features at the microsystem level. The most distal level of the model is termed the *macrosystem*, which includes the general beliefs, values, customs, and laws of the society in which the child is embedded. At the base of Bronfenbrenner's model is the fifth, and final, system – the *chronosystem*. In the original model, the term chronosystem was used to refer to the temporal dimension of development.

Over time, Bronfenbrenner's understandings of development, and therefore the design of his model changed and evolved. Prior to his death, he co-authored a chapter in Lerner's *Handbook of Child Psychology: Volume 1*, in which he discussed its four

defining properties. Two of these – context and time, are integral to the chronosystem, while the other two – process and person, are not (Bronfenbrenner & Morris, 2006). He also refined his definition of process, re-naming it *proximal processes*, which he now argued were the primary mechanisms driving human development, and varied as a function of the person, context, and time periods (Bronfenbrenner & Morris, 2006). The notion of person (or individual) was also re-conceptualised. He outlined three different characteristics of people that can influence their future development – dispositions, bioecological resources or ability, experience, knowledge and skills, and demand characteristics (Bronfenbrenner & Morris, 2006), and argued that individuals who acquired experiences and developed knowledge and skills were better able to realise healthy developmental outcomes.

Bronfenbrenner’s model posits that to realise optimal social health, children must be offered opportunities to develop social skills, healthy social relationships, and strong social networks, through positive, bi-directional relationships with features in their proximal environment (school, family, and peers). By implication, the model also posits that without such opportunities, social health may be unattainable. This may certainly be the case for children in poverty because of the inherent “hecticness, instability, and/or chaos that exist in their microsystems” (Bronfenbrenner & Morris, 2006, p. 796). Improved social health outcomes for these children therefore, will likely require targeted intervention at the microsystemic level – programmes that will promote relationship management and social skills, and support children to develop the competencies that will enhance resilience and promote healthier social attitudes and behaviours. In short, a positive developmental approach.

## **2.6 Positive Youth Development Theory**

In the latter half of the 20<sup>th</sup> century, interest in adolescent development and the concept of resilience inspired researchers, practitioners, and policy makers to consider how young people might not only survive, but thrive, in the face of challenges, obstacles, and deficits (Benson, Scales, Hamilton, Sesma Jr., Hong, & Roehlkepartain, 2006). It was a significant shift away from the deficit orientation that had historically pervaded health practice and scholarship, and the thinking, philosophies, and strategies that emerged, would later coalesce and be known as positive youth development theory.

Positive youth development theory is based on two assumptions – that every child has the potential for health, and that all children are capable of developing positively (Lerner, et al., 2002). It is premised on the idea that every child has unique talents, interests, and skills that merit strengthening and nurturing (Benson et al., 2006; Peterson, 2004), and it presupposes an image of children as eager to explore, gain competence, and acquire the capacity to contribute to their world (Peterson, 2004). In essence, it sees children as “resources to be developed” (Lerner et al., 2005, p. 11) rather than as “problems to be managed” (Roth & Brooks-Gunn, 2003, p. 97).

Benson et al. (2006) note that there is widespread agreement about the philosophies underpinning the approach, but as yet little consensus on how to define the term itself. In the absence of an agreed definition, they direct focus to the theory’s core constructs – the developmental context in which the child exists, the child’s specific developmental strengths, their capacity to engage with all developmental contexts, the promotion of overall thriving, and the desire to reduce high-risk behavior. The theory posits that positive development results from sustained, reciprocal interactions between individuals, and the contexts in which they occur, and draws on bioecological understandings of human development to explain how such interactions lead to positive developmental outcomes (Catterall, 1998; Crosnoe, Erickson, & Dornbusch, 2002; Leffert, Benson, Scales, Sharma, Drake, & Blyth, 1998). Positive outcomes occur, it argues, when children acquire internal and external skills or assets (Catterall, 1998). This notion of asset acquisition is at the core of Benson’s 40 developmental assets framework.

### **2.6.1 Developmental assets framework.**

The developmental assets framework was created by Peter Benson (and colleagues) at The Search Institute in Minneapolis in 1990 (Benson, 1990, 1997, 2006; Search Institute, 2011a), and was based on research in the positive development, resiliency, and disease prevention domains (Search Institute, 2011b). It was subsequently refined and expanded in 1995. In 1990, improving health outcomes for American children was a key focus of the Search Institute, and the development of assets in childhood was seen as critical to that aspiration (Benson & Leffert, 2001; Scales, 1999). In the original theoretical model and framework that followed, 30 developmental assets were identified for healthy, positive development (Search Institute, 2011a). The team then

recruited a sample of 500,000 children from 600 American communities to test and evaluate its sufficiency. Revisions followed, and in 1996, 10 new assets were added (Scales, 1999). Benson (2006) described the assets as the:

key relationships, opportunities, values, skills and self-perceptions that help young people limit their engagements in high-risk behaviours, enjoy resilience in the face of adversity, and thrive. (p. 2)

The now 40 assets were categorised as *internal assets* or *external assets*, and then divided into four subsections (Low & Jose, 2011). The subsections in the external assets category were – support-related assets (Assets 1-6), empowerment-related assets (7-10), boundaries and expectations-related assets (11-16), and constructive use of time-related assets (17-20). The subsections in the internal assets category were – commitment to learning-related assets (21-25), positive value-related assets (26-31), social competency-related assets (32-36), and positive identity-related assets (37-40). The framework included seven assets that were specifically related to social skill development, and healthy social attitude and behavior acquisition – caring, equality and social justice, interpersonal confidence, cultural competence, peaceful conflict resolution, planning and decision making, and resistance skills.

In 1998, Leffert et al. evaluated the content and construct validity of the framework using a factor analysis technique, and discovered that the most appropriate 40 developmental assets model had a 16-factor solution. This solution accounted for 50% of variance and each of the 16 factors fell into (at least) one of eight asset categories. The following year, Scales tested for reliability and validity, and found that 19 assets had alpha reliabilities of Cronbach's  $\alpha = .60 - .80$ , eight assets had alpha reliabilities lower than  $\alpha = .60$ , and 13 assets were measured using a single item (Scales, 1999). In 1999, following further research, evaluation, and consultation (Scales, Sesma, & Bolstrom, 2004; Search Institute, 2011c; VanderVen, 2008), the framework remodeled in 1995 (now known as the '40 assets developmental framework') was deemed too generic, and was replaced with four, age-specific frameworks (Search Institute, 2011b).

In the early 1990s, two other tools were developed to measure the expression of assets in American youth (Benson, Leffert, Scales, & Blyth, 1998; Leffert et al., 1998) – the Profile of Student Life: Attitudes and Behaviours (PSL: AB) survey, and the Developmental Assets Profile (DAP). Both were based on the 40 developmental assets framework. The PSL: AB is the longer of the two. It measures all 40 assets and includes items related to student thriving, development, attitudes and risk-taking (antisocial) behaviours. It was designed and tested with samples of children in Grade 6 to Grade 12, and has demonstrated adequate reliability and validity (Benson et al., 1998; Search Institute, 2011b). The DAP, is a shorter, online version of the PSL: AB. It, too, measures student asset accumulation but is less popular because it does not include behavioural and/or attitudinal items. The Search Institute (2013) reports that the DAP has demonstrated reliability and validity with two large samples of children (1,300 children in Grades 6 to 12 in Minnesota, and 1,100 children in Grades 6 to 8 in Oregon), and that it has good internal consistency (average  $r = .81$ ), test-retest reliability (average  $r = .79$ ), and adequate internal reliability in cross-cultural studies of asset accumulation (Scales, 2011).

The PSL: AB is the now the most widely used tool for measuring positive youth developmental outcomes in children in the United States (Search Institute, 2011b). To date, it has been used with over 1,000,000 children in approximately 1,000 communities (Scales, Benson, Leffert, & Blyth, 2000). Significantly, the results it has yielded suggest that asset accumulation is affected by gender, age, and socioeconomic status. Benson, Scales, Leffert, and Roehlkepatain (1999) in their study of 99,462 adolescents, for example, found that boys reported three fewer assets, on average, than girls the same age, and that the number of assets decreased as children got older. Overall only 8% of the sample had 31 of the 40 assets required to be deemed *asset-rich*, 30% reported 21-30 assets, 42% reported 11-20 assets, and 20% reported 0-10 assets. The most commonly reported assets amongst participants in this study were related to children's families and religious communities – 60-74% of all participants reported that they had a positive view of their personal future, family support, and adequate time spent with others in their religious community (Benson et al., 1999). The least commonly reported assets were related to the school environment – only 25% of students reported membership in a caring school (Benson et al., 1999), a finding that suggests the need for additional work to ensure that children feel safe,

connected, and cared for in the school setting. In a second related study assets were measured in children from a low-decile school in Minnesota (Roehlkepartain, Benson, & Sesma, 2003). In this longitudinal study, the researchers documented changes in asset accumulation over time. At baseline they reported that the average student experienced half of all possible assets (20.4 of 40.0). By the end of the four years, however, the average student reported an increase in the number of developmental assets (Roehlkepartain et al., 2003). When the data were segregated by age, the number of developmental assets appeared to decline sharply from Grade 6 to Grade 8 (from 25.0 to 20.7 assets for females and from 22.2 - 17.6 assets in males), remained low from Grades 8 to 12, and then rose again (Roehlkepartain et al., 2003). This suggests that in low-decile environments students may struggle to acquire the assets that they require to experience positive developmental outcomes, particularly during the intermediate and adolescent years. Unfortunately, little is currently known about patterns of asset acquisition for children in low-decile primary schools.

To date, few studies have looked specifically at the relationship between asset development and health for children in low-decile primary and intermediate schools. Those that have, generally report a positive linear relationship – as adolescents accrued more assets, they were more likely to form attitudes and engage in behaviours that promoted superior health outcomes. A study in 1997 of adolescents in the mid-West of the United States, reported that asset-rich children were significantly less likely to suffer from depression or to engage in violent or high-risk behaviours (such as excessive drinking or smoking), than their *asset-poor* or *asset-average* peers (Benson et al., 1998; Leffert et al., 1998; Scales et al., 2000), findings that were consistent across gender, age, and ethnicity. A second, related study reported that asset-rich youth were also more likely to succeed at school and to be physically active (Benson et al., 1999). Even fewer studies have explored the relationship between asset accumulation and social health outcomes. In the one that has (Benson et al., 1999), the researchers did not elaborate on the impact of asset accumulation on the social attitudes, behaviours, or health of the children, reporting simply that vulnerable children who accumulated developmental assets were more likely to maintain a state of good overall health and value diversity than those without developmental assets.

### 2.6.2 Lerner's 5C's theory of positive youth development.

The 5C's theory of positive youth development evolved in the late 90s, and was largely the work of by Dr. Richard Lerner. The 5C's theory is based on many of the underlying principles of Benson's frameworks, but whereas Benson lists the assets necessary for positive development, Lerner presents a picture of an actualised positive youth, a child who has accrued assets, has moral character, and who is competent, confident, connected, and caring/compassionate – a child who is, in effect, *flourishing* (Lerner et al., 2010). Unlike its predecessor, the 5C's model is also based on developmental systems theory, and therefore is also influenced by the work of comparative psychologists and biologists from the early 20<sup>th</sup> century. These scientists explored the plasticity of developmental processes that arose from interactions between individuals and their environment (Lerner et al., 2010).

Lerner et al. (2010) contend that the 5C's theory can be used to explain how individuals experience either positive (adaptive) or negative (maladaptive) development. They posit that youth development is the product of a systematic, but dynamic, alignment between the developing individual and their changing environment, a concept they refer to as *bi-directional, individual-context relations* (Lerner et al., 2010). Positive (or adaptive) development, they argue, occurs when these individual-context interactions are advantageous for both parties (Lerner, Lerner et al., 2005). This is possible at all ages, but particularly so during childhood when children possess a unique plasticity, or ability to undergo systemic change (Lerner et al., 2010). Lerner used the term *thriving* to describe the manifestation of attributes that mark a healthy, positive, flourishing youth, and there were five such attributes in his original model – the 5C's: Competence, Confidence, Connection, Character, and Caring/Compassion. Competence, he argued, refers to domain-specific abilities within four areas – social, cognitive, academic, and vocational competence, and in this model, social competence referred specifically to interpersonal skills. Confidence related to self-worth, self-efficacy, and global self-regard. Connection referred to the bonds that exist between the individual and their peers, family, school, and community, and those bonds were expected to be bidirectional, positive, and mutually beneficial. Caring/Compassion referred to the individual's ability to demonstrate

sympathy and empathy, and character referred to the ability to demonstrate respect and integrity, and moral strength (Lerner et al., 2010).

Over time, the 5C's model was modified, and a sixth 'C' – Contribution, was added to the framework (Bowers et al., 2010). Children who possessed all six C's were considered more likely to thrive and be on a positive life trajectory, and less likely to engage in high-risk behaviours (such as substance abuse and delinquency), or to suffer from depression (Lerner et al., 2005; Phelps, Balsano, Fay, Peltz, Zimmerman, & Lerner, 2007). Significant empirical work has been undertaken by American researchers eager to establish (or not) the veracity of 5C's theory and its assertions (Gestsdottir & Lerner, 2007). The most notable study, which began in 2004 and is ongoing, is the 4-H Study of Youth Development, a longitudinal assessment of positive development amongst children from Grade 5 to Grade 12 (Bowers et al., 2010). This study involves 6,120 children (59% female) and 3,084 parents from 43 states (Lerner, Lerner et al., 2005), from a diverse range of ethnic, racial, religious and/or socioeconomic backgrounds (Lerner, 2004; 2007; Lerner et al., 2005; Lerner et al., 2010). In this study, a tool for measuring positive youth development was created by the research team. The tool comprised five scales, with each scale capturing development in one of the five 'C' domains. The items for each scale were derived from different measures. Items for the first scale, Competence, and for a subscale of Confidence (Self-Worth), were drawn from the Self-Perception Profile for Children (Harter, 1982). Items for the second Confidence subscale (Positive Identity), and the Connection to School, Connection to Family, Connection to Community subscales, and the Character scale, were drawn from the Search Institute's PSL: AB survey (Benson et al., 1998). The Connection to Peers subscale was measured using items from the Teen Assessment Project Survey Question Bank (Small & Rodgers, 1995), and items for the Caring/Compassion scale were drawn from the Eisenberg Sympathy Scale (Eisenberg, Fabes, Murphy, Karbon, Smith, & Maszk, 1996) and the Interpersonal Reactivity Index (Davis, 1980). The scale reliability of the positive youth development tool and structural validity of the theoretical model were assessed by Phelps, Zimmerman, Warren, Jellicic, von Eye, and Lerner (2009), who undertook a three-year, longitudinal study in which they administered the tool to a sample of Grade 5 ( $n = 1,722$ ), Grade 6 ( $n = 1,967$ ), and Grade 7 ( $n = 778$ ) children in the United States. The tool was administered once a year for three years, and the results

indicated that all five positive youth development scales had adequate internal reliability ( $.62 > r < .89$ ). The results from this study supported those of earlier research by Lerner, Lerner et al. (2005) who used a CFA technique to demonstrate the empirical validity of the 5C's and the positive youth development construct with a sample of Grade 5 children in the United States. They reported that the final structural positive youth development model had five, first-order latent factors (one for each of the five Cs) and one second-order latent factor, positive youth development ( $X^2 = 552, df = 134, p < .01; RMSEA = .043; CFI = .99$ ).

In a study by Phelps, Zimmerman, Warren, Jelicic, von Eye, and Lerner (2009), LISREL 8.8 was used to re-estimate Lerner's model and to estimate structural positive youth development models for a sample of Grade 5, 6 and 7 students. The results of the CFA procedures indicated that the second-order positive youth development model had adequate fit for children in Grade 5 ( $X^2 = 1,185.70, df = 89, p < .001; RMSEA = .065, 90\% CI [.062, .069]; CFI = .10$ ), Grade 6 ( $X^2 = 1,926.85, df = 136, p < .001; RMSEA = .068, 90\% CI [.065, .070]; CFI = .10$ ) and Grade 7 ( $X^2 = 2,234.86, df = 107, p < .001; RMSEA = .083, 90\% CI [.080, .086]; CFI = .10$ ). Further, the model explained 21% (Caring) to 85% (Confidence) of the factor variance in Grade 5, 36% (Competence) to 78% (Confidence) for Grade 6, and 36% (Competence) to 90% (Connection) for Grade 7 (Phelps et al., 2009). When the models were compared to assess structural co-variance over time, the correlations across positive youth development subscales were high for the Grade 5 to Grade 6 ( $r = .88$ ) and Grade 6 to Grade 7 ( $r = .89$ ) samples. They were also high amongst a sample of Grade 8 to Grade 10 students. Bowers et al. (2010) investigated the structural invariance of a revised, second-order positive youth development model with adolescents and found that it demonstrated strict configural, metric, scalar, and factorial invariance ( $\Delta CFI < .10$ ).

These results provide strong support for the validity and invariance of the 5C's model. They suggest that the model can be used to assess positive developmental change in children of different ages, year levels, and geographic locations. As a result of this psychometric work, a number of researchers have embraced the concept of positive youth development and sought opportunities to investigate the 5C's in primary school aged children (Bowers et al., 2010; Heck & Subramaniam, 2009; Lerner, Lerner et al.,

2005; Phelps et al., 2009). In most cases, they have used the positive youth development tool described above to explore the relationship between the C's, student health, and behaviour.

In one of the most widely cited 4-H studies, Jellicic et al. (2007) investigated the relationship between the 5C's, depression, and risk taking behavior in a sample of Grade 5 and Grade 6 American children, reporting that higher scores for positive youth development subscales in Grade 5 predicted lower incidence of risk behaviours and depression in Grade 6. This finding led the research team to suggest that individuals who accumulated more 'C's', were less likely to experience risky or antisocial behaviours as they aged (Jellicic et al., 2007). In a related study, Bowers et al. (2011) investigated the predictive effect of asset accumulation on the incidence of antisocial behaviour in a sample of Grade 5 to Grade 11 American children ( $n = 626$ ), and reported that children were more likely to demonstrate antisocial behaviour when they attended schools devoid of asset building opportunities, activities, and social support (Bowers et al., 2011). Naturally they concluded that schools should adopt programmes that foster social skills, social networks, social bonding, and social cohesion, and that such programmes should use a positive youth development approach.

## **2.7 Positive Youth Development Programmes**

Advocates of positive youth development theory argue that positive youth programmes can enable children to become civically responsible individuals who actively contribute to their communities, and support them to be equitable, socially just, and democratic (Lerner et al., 2002). Peterson (2004) reports that there are a number of positive youth development programmes currently operating throughout the world, and while most are fully aligned with the tenets of positive youth development (aimed at strengthening existing skills and enhancing health and wellbeing), a few are more prevention-based in approach (aimed at reducing risk-taking or health-compromising behaviours). Positive youth development programmes share a belief that positive outcomes follow when children are listened to and treated respectfully, and when there is an expectation that they will engage positively with, and invest in, their communities (Peterson, 2004). Those with a specific social development focus aim to promote bonding, resilience, social competence,

behavioural competence, moral competence, and prosocial norms (Catalano et al., 2004). These programmes are typically delivered in-class or after school, and may include interventions in a range of settings – schools, homes, churches, worksites, etc. (Peterson, 2004).

In 2004, Catalano et al. (2004) conducted an extensive review of the positive youth development literature, focusing in particular on the impact that participation in positive youth development programmes had on children's social development and behaviour. Of the 161 programmes reviewed, 19 demonstrated efficacies in improving interpersonal skills, the quality of peer and adult relationships, and enhancing problem solving amongst participants, while a further 24 reported a significant reduction in the incidence of aggressive and violent behavior. In many of these studies, the programmes were offered in a single setting (typically a school or a community facility), rather than across multiple settings. Catalano et al. reported that participation in such programmes was associated with greater self-control, assertiveness, and adaptive coping in peer-pressure situations, while participation in programmes offered only at schools was associated with better personal health-management attitudes and knowledge, better health practices, greater assertiveness, sociability, problem solving and frustration tolerance, increased acceptance of prosocial norms, higher capacity for managing personal reactions and behaviour, greater self-efficacy, and increased empathy.

Similar results have been reported by researchers in the United States interested in exploring the impact of participation in after-school programmes on student social attitudes and behaviours. According to Shernoff (2010), middle school aged students in the Midwestern United States ( $n = 4970$ ) reported higher social competence than non-participants after a year of after-school programme participation. In this study social competence was a composite measure of student social skills, student engagement in the school classroom and positive/negative affect (Shernoff, 2010). Improved social skills and emotional self-understanding were also reported by Larson and Brown (2007) in their sample of 20 adolescents participating in an after-school theatre programme in the United States. For others, improved social attitudes and behaviours were reported amongst participants of school-based interventions. Two of the most commonly cited examples are: The Good Behaviour Game (Barrish,

Saunders, & Wolf, 1969) and the Seattle Social Development Project (Lonczak, Abbott, Hawkins, Kosterman, & Catalano, 2002). The Good Behaviour Game was initially designed as a classroom-based intervention for reducing disruptive, or antisocial, behaviour. In most cases, teachers divide their class into teams, establish rules or 'winning criteria' (Keenan, Moore, & Dillenburger, 2000, p. 27) and reward prosocial behaviour at the team level with confectionary or stationary items. Participation in The Good Behaviour Game has been associated with a marked decrease in observed rates of antisocial behaviour amongst i) boys with emotional and behavioural difficulties in Northern Ireland (Keenan, Moore, & Dillenburger, 2000), ii) kindergarten students ( $n = 81$ ) in west Texas (Donaldson, Wiskow, & Soto, 2015); iii) elementary students in Sudan (Saigh & Umar, 1983) and the Netherlands (van Lier, Muthen, van der Sar, & Crinjen, 2004), and iv) middle school-aged students in the United States (Hendricks Brown, Kellman, Ialongo, Poduska, & Ford, 2007). In each of these studies, the positive effects of participation were limited to the implementation phase of the programme and were not sustained over time. This is not the case for the Seattle Social Development Project. In fact, participation in the Seattle Social Development Project has been associated with a reduction in violent behaviour (Herrenkohl, Jungeun Lee, & Hawkins, 2012), a decrease in incidence of unsafe sexual behaviour (Lonczak et al., 2002) and improved emotional and mental competence (Hawkins, Kosterman, Catalano, Hill, & Abbott, 2005) amongst programme participants later in life. Of the two programmes, the Seattle Social Development Project is the most closely aligned with positive youth development theory. It is based on a social development model that focuses on enhancing the socialization processes of students in Grades 1 to 6 in 18 Seattle public elementary schools. It began in 1985 and assessments of student behaviour are undertaken every three years (Herrenkohl, Jungeun, & Hawkins, 2012).

In New Zealand, positive youth development programmes, similar to the Seattle Social Development Project, are typically offered in single settings (usually schools). There are currently three programmes that actively promote social health and development in the schooling context – Cool Bananas, Cool Schools, and Kiwi Can. Cool Bananas promotes positive values and character development in Christian primary schools in the Tauranga region (Cool Bananas, 2012), while Cool Schools is a national initiative, administered by the New Zealand Peace Foundation, and

promotes the development of peaceful conflict resolution skills (New Zealand Peace Foundation, 2014). Kiwi Can is far broader in scope and reach than either Cool Bananas or Cool Schools. It is a stand-alone, externally provided social health promotion programme that operates on a weekly basis in 64 low-decile primary and intermediate schools throughout New Zealand. It is delivered to all the children in all the schools that participate. Kiwi Can lessons focus on the development of Lerner's 5C's – Competence, Confidence, Connection, Character, and Caring/Compassion, but with a social health emphasis. Its' aim is to promote social health in vulnerable children, and it is intended to complement the social health classes that are part of the compulsory health education curriculum (Ahmed, 2010).

## **2.8 Health Education Curriculum in New Zealand**

In the New Zealand school curriculum, health education is currently positioned alongside physical education and home economics, and together the three subjects comprise a learning area called *Health and Physical Education* (HPE). HPE is compulsory in Years 0 to 10, or for all children aged 5 to 15 years (MOE, 2007), but teachers enjoy considerable freedom around programme delivery. Penney (2008), and Smith and Philpot (2011) believe that, in many schools, that freedom has been problematic, and that the delivery of core HPE subjects is often inequitable. Sinkinson (2011) concurs, arguing that the delivery of physical education has frequently been favoured over the delivery of health education.

This was not always the case however – when health education was first introduced to New Zealand schools in the early 20<sup>th</sup> century, it was taught as a stand-alone subject (Barlow, 1990; New Zealand Education Department [NZED], 1928, 1935). Then, it was informed by biomedical notions of health and disease, and by the rules of good behaviour (NZED, 1928, 1935). Teachers offered lessons in cleanliness, posture, suitable foods, first aid, thrift, and the need for fresh air (Ewing, 1970) – all underpinned by a strong moralising discourse and the expectation that teachers would “impress on pupils that health is our greatest gift, and that healthful living is not only a duty to the individual himself, but also a preparation for future efficient citizenship” (NZED, 1928, p. 202; 1935, p. 204). As understandings about health changed however, so too did the way that health education was conceptualised and delivered. Traditional approaches had failed to reduce risk taking and unhealthy behaviour in

children, and health was now more popularly viewed as inter-connected, inter-dependent, and multi-disciplined (Bronfenbrenner, 1979; Durie, 1999). In schools, disease prevention approaches gave way to health promotion (Sinkinson, 2011), and an emphasis on hauora (the Māori philosophy of wellbeing), attitudes and values, and socioecological perspectives (MOE, 2007).

In 1999, the HPE curriculum was released. This document was shaped by two powerful forces – a neoliberal political agenda, and reaction to the 1994 UNICEF report entitled *The Progress of Nations*, in which New Zealand's rates of child suicide, sexually transmitted diseases, and alcohol related problems compared poorly to those of other OECD countries (Grant, 1994). Ovens and Cassidy (2009) state that The Progress of Nations report sent shock waves through New Zealand, and created a climate of crisis within the health sector. The first draft of the new curriculum was completed only weeks after its release, following a Ministry directive (Ovens & Cassidy, 2009; Sinkinson & Burrows, 2011), and for the first time, mental health and sexuality education were identified as 'key' areas of learning, alongside food and nutrition, body care and physical safety, sport studies, outdoor education, and physical activity (MOE, 1999), and remain as priorities in the current curriculum document.

The current HPE curriculum is one of eight learning areas in New Zealand primary schools (MOE, 2007). It is composed of four strands (physical health, mental health, emotional health, and social health), seven Key Learning Areas (KLAs), and has 11 Achievement Objectives (AOs). Three of the four strands, four of the seven KLAs, and nine of the 11 AOs relate directly to learning in the health education area, and the entire document is based on the same four principles inherent in the original – hauora, attitudes and values, health promotion, and a socioecological perspective (MOE, 2007). Within the social health dimension specifically, children are expected to develop healthy relationship skills, the ability to problem solve and resolve conflicts, socioecological understandings, and the confidence to act in the interests of their own wellbeing, and the wellbeing of others and society as a whole.

Overall, it is an ambitious curriculum, one that requires teachers with a passion for health and specialized health knowledge to deliver effectively. McGee et al. (2004), in the only review of HPE delivery to date, report that this is not always (or often) the case. Their study involving 1,716 primary and secondary school teachers throughout

the country, reports that in some schools, health and physical education were only taught sporadically, and often by teachers with no specialised health education training.

Such results are not unique to New Zealand. Researchers in Sweden have reported that health education is a marginalised area of the compulsory curriculum and is delivered infrequently, if at all, in state schools across the country (Bergstrom, 1988; Karlsson & Rastam, 1995). In the Auvergne and Nord Pas de Calais regions of France, a survey of 625 Year 1 to Year 5 teachers revealed that only 34% of participants delivered a comprehensive health education programme – the rest provided only a basic programme, or no programme at all (Jourdan, Pommier, & Quidu, 2010). In the Acacia prefecture in Greece, researchers report that health education delivery in primary schools was also sporadic, and hindered by specific, personal barriers – the nature and strength of which differed between teachers as a function of experience. For teachers with over 15 years experience, a lack of specialised training, limited experience teaching health education, and negative attitudes toward the topic were cited as issues, while less experienced teachers struggled with the negative attitudes of the children in health education classes (Cholevas & Loucaides, 2012). These results corroborate those from previous studies conducted in Scotland (Buston, Wight, Hart, & Scott, 2002), Canada (Dwyer, Allison, Barrera, Hansen, Goldenberg, & Boutilier, 2003), and the United States (Kincaid, Childs, Blasé, & Wallace, 2007; Lytle, Ward, Nader, Pederson, & Williston, 2003; Pateman, Grunbaum, & Kann, 1995). Other researchers indicate that teacher beliefs can have a strong effect on practice, particularly in areas that are highly sensitive, such as health education. Howard-Barr, Rienzo, Pigg, and James (2005), in a study on teacher beliefs and the delivery of sexuality education, reported that teachers who did not believe in the importance of educating children about sexuality, were significantly less likely to deliver programming in the area than those who did. Iyer and Aggleton (2013) report similar results in a study involving teachers and sexuality education at a private, co-educational secondary school in Uganda. Here religious convictions about abstinence meant that teachers often chose to deliver health education lessons in keeping with their own beliefs, rather than following the school curriculum.

In New Zealand (and overseas), efforts have been made to overcome barriers to health education (lack of time, lack of resourcing and the negative attitudes and beliefs of children and teachers) by hiring external providers of health and physical education programmes in schools (Dyson et al., 2011; Morgan & Hansen, 2007, 2008; Petrie et al., 2014; Williams, Hay, & Macdonald, 2011). A recent stocktake of externally provided health and physical education programmes (Petrie et al., 2014) suggests that dozens of providers are currently working in New Zealand primary schools, but that virtually nothing is known about the implementation or impact of these programmes (Dyson et al., 2011). In Australia, some empirical work has been undertaken in this area, but it has generally been narrow in scope, focused exclusively on physical education in primary schools (Griggs, 2008, 2010; Lavin, Swindlehurst, & Foster, 2008; Lynch, 2007), and supported by data that were partial or fragmented (Williams et al., 2011). This has limited the validity, reliability, and generalisability of the results.

## **2.9 Research Questions**

At present there are a large number of vulnerable children living in New Zealand. Many of these children are living in poverty and some are forced to move with their families on a regular basis in search of affordable housing and employment. These transient children are at increased risk of adverse social health and behavioural outcomes. The national government has recognised the need to support these children and has called upon educators (and others) to provide opportunities within the schooling system to promote social health and wellbeing. Social health programming is currently offered through the provision of health education lessons. The lessons are guided by the national HPE curriculum document but they do not appear to be effective in addressing the myriad of social health issues that face vulnerable children in low-decile communities. Therefore, there may be a need to introduce alternative social health promotion programming in low-decile schools. Outsourced social health programmes, particularly those based on positive youth development theory, may be one means of building resilience and social health, however efficacy trials are needed to ensure that such programmes are addressing the social health needs of children.

In this thesis, the implementation and impact of an outsourced, school-based social health promotion programme on the social attitudes and behaviours of vulnerable children in low-decile schools around New Zealand will be conducted. This thesis

will also evaluate the implementation of the current health education curriculum in New Zealand primary and intermediate schools. The empirical work is guided by 19 research questions.

For Study 1 (the delivery of the health education curriculum in New Zealand primary and intermediate schools), the research questions focus on three main constructs – frequency (of delivery), barriers (to delivery), and beliefs (about delivery). The questions were designed to gain a baseline understanding of general practice in health education, and more specifically, to explore factors affecting social health education delivery. For Study 2 (the implementation of the Kiwi Can curriculum), the research questions focus on the quality and fidelity of programme delivery, and participant responsiveness. For Study 3 (the outcome evaluation of Kiwi Can), the research questions focus on the impact that participation had on the social attitudes and behaviours of vulnerable children in low-decile primary and intermediate schools. The research questions were written to capture change in individual attitude and behaviour scores amongst programme participants and the control group over the course of a single school year. The specific research questions for each study are outlined below.

**Study 1: Health education implementation in primary and intermediate schools.**

1. How frequently is health education, as a general area of study, taught in primary and intermediate schools around New Zealand?
  
2. How frequently do primary and intermediate school teachers deliver lessons across the four health dimensions, four KLAs, and 11 AOs, contained in the health education curriculum?
  - 2(i) How frequently are social health-related lessons delivered in primary and intermediate schools?
  
3. What impact do teacher beliefs about the importance of teaching a health dimension have on the delivery of lessons across the four health dimensions, four KLAs, and 11 AOs?

- 3(i) What impact do teacher beliefs about the importance of teaching a health dimension have on the delivery of social health-related lessons?
4. What impact do teacher beliefs about the benefit of teaching a health dimension have on the delivery of lessons across the four health dimensions, four KLAs, and 11 AOs?
- 4(i) What impact do teacher beliefs about the benefit of teaching a health dimension have on the delivery of social health-related lessons?
5. What are the most frequently cited barriers to delivery of the health education curriculum in primary and intermediate schools around New Zealand?
6. Are there statistically significant differences in self-reported delivery of health education as a function of school position, years teaching health education, and/or year level taught?

**Study 2: The implementation evaluation of Kiwi Can.**

1. According to Year 4-6 student participants, how well was the Kiwi Can programme implemented in low-decile primary schools across New Zealand? (Quality)
- 1(i) To what degree did Year 4-6 student participants report learning new skills and values? (Quality)
- 1(ii) To what degree did Year 4-6 student participants report satisfaction with the Kiwi Can programme structure, organisation, and leadership? (Quality)
- 1(iii) To what degree did Year 4-6 student participants report enjoying the Kiwi Can programme? (Participant responsiveness)
2. How frequently did students report learning about prosocial values and practicing social skills in their Kiwi Can classroom during the 2012 school year? (Fidelity)
3. Did participant responses differ by school condition, year level, sex, and/or ethnicity?

### **Study 3: The outcome evaluation of Kiwi Can.**

1. What impact does participation in Kiwi Can have on the self-reported incidence of individual prosocial attitudes and behaviours, amongst vulnerable students in low-decile primary and intermediate schools?

1(i) Are there statistically significant differences in the measures of individual prosocial attitudes and behaviours as a function of school condition (i.e., new Kiwi Can, experienced Kiwi Can, and/or control school), sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools at baseline?

1(ii) Are there statistically significant differences in measures of individual prosocial attitudes and behaviours as a function of school condition, sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools over the course of one academic year?

2. What impact does participation in Kiwi Can have on the self-reported incidence of prosocial attitudes and behaviours at the classroom level, amongst vulnerable students in low-decile primary and intermediate schools?

2(i) Are there statistically significant differences in the measures of classroom climate as a function of school condition, sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools at baseline?

2(ii) Are there statistically significant differences in the measures of classroom climate as a function of school condition, sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools over the course of one academic year?

3. What impact does school-level transience have on self-reported incidence of prosocial attitudes and behaviours at baseline and across time for children in low-decile primary and intermediate schools across New Zealand?

4. What impact does school-level transience have on self-reported measures of classroom climate at baseline and across time for children in low-decile primary and intermediate schools across New Zealand?
5. Are there statistically significant differences in measures of individual prosocial attitudes and behaviours or classroom climate for children as a function of school-level transience and school condition?
6. What is the magnitude and direction of the interaction between school condition and transience level on self-reported incidence of individual prosocial attitudes and behaviours, and measures of classroom climate for children at baseline and across time?

## Chapter 3: Theoretical Methodology

This chapter presents an overview and rationale for the methods used in the three studies outlined in the introductory chapter, as well as an overview of the theory underpinning the programme evaluation component. It is divided into two sections. Section one (3.1) presents a brief history of programme evaluation, including a description of the theoretical model that served as the guiding framework for the evaluation component. Section two (3.2) presents a description of the quantitative, survey-style research methodologies used, and in the subsections, literature is provided to support the choices made about experimental design and sampling, the selection of data preparation methods, and the selection of data analysis methods.

### 3.1 Programme Evaluation Theory

This thesis specifically evaluates the Kiwi Can intervention, but in general terms, it is also an evaluation of the New Zealand health education curriculum.

The field of programme evaluation first emerged in the United States in the 1960s (Patton, 1997), and in the decades that have followed, researchers and scholars have developed systematic processes for determining the legitimacy and worthiness of programmes (Patton, 1997; Shadish, Cook, & Leviton, 1991). As most evaluators considered that the truth about a phenomenon could only be obtained by measuring and manipulating variables and making direct, objective observations of cause and effect (Trochim & Donnelly, 2008), quantitative-experimental paradigms have traditionally been employed (Hall & Hall, 2004; House, 1993; Mark, 2003; Patton, 1997, 2002), including the randomised control trial design (Campbell & Stanley, 1963; Cook, 2000; Cook & Campbell, 1979; Lipsey & Cordray, 2000; Rossi & Freeman, 1989; Shadish, Cook, & Campbell, 2002; Stufflebeam, 2001; Trochim & Donnelly, 2008; Weiss, 1998). In most cases, evaluators used the randomised control trial design to assess whether programme objectives were being met (outcome evaluation), to compare programmes to one another (comparative assessments), and to determine the cost-effectiveness of the initiatives (cost-benefit analyses). These practices are now collectively known as *summative evaluation* techniques (Hall & Hall, 2004; Patton, 1997, 2002).

Criticism of the quantitative-experimental approach (Hall & Hall, 2004; House, 1993; Patton, 1997; Weiss & Rien, 1969, 1970) focused on four concerns: (i) that many evaluations were failing to show evidence of an effect (House, 1993); (ii) that the truth behind complex social programming could not be elucidated using a small number of measurable performance indicators (Patton, 1997); (iii) that programme stakeholders were being alienated during the evaluation process (House, 1993; Patton, 1997)]; and (iv) that insights into how or why a programme operated were not provided (Chen, 1990; Patton, 1997; Stufflebeam, 2001; Weiss & Rein, 1969, 1970). These shortcomings, it was argued, limited the effectiveness of evaluations, and meant that providers and programme funders were not able to determine how to improve programme design and delivery effectively (Astbury & Leeuw, 2010; Campbell, 1970; Chen, 1990; Donaldson, 2003, 2007), and quickly led to what Patton (1997) called a *utilisation crisis* and significant change.

The new approaches that followed, favoured the inclusion of stakeholders and the investigation of programme processes (Hall & Hall, 2004; House, 1993; Mark, 2003; Patton, 1997; Weiss, 1998). These process-driven, stakeholder-orientated evaluations employed qualitative-naturalistic methodologies that were markedly different from the traditional quantitative-positivist methods (Weiss, 1998). They were more holistic than standardised quantitative methods, provided in-depth accounts of the programme experience (Patton, 1997, 2002), included options that ranged from liberal (empowerment evaluation) to conservative (stakeholder evaluation), and proved very popular amongst programme evaluators (Weiss, 1998).

The new methodologies also had their critics however, and concerns focused on rigour, credibility, and the priority given to the interests of the stakeholders in the process-driven approach (Stufflebeam, 1994). A clear division soon emerged between those who favoured the quantitative-experimental approach, and those who favoured the qualitative-naturalistic design. This *paradigm war* is ongoing (Denzin & Lincoln, 2005; Newman & Hitchcock, 2011; Smith & Hodkinson, 2005; Weiss, 1998), though in many cases, decisions about design are driven by contextual factors (and not personal paradigmatic affiliation) and many evaluators now choose to incorporate elements of both approaches (Chen, 1994; Donaldson, 2003, 2007; Mark, 2003; Shadish, Cook, & Leviton, 1991). Those who choose to engage in *contextually-*

*driven or contingency-based* evaluations (Shadish, Cook, & Leviton, 1991) use a variety of frameworks for operationalising their practice (Brazil, 1999; Cousins & Whitmore, 1998; Nastasi & Hitchcock, 2009; Patton, 1994, 1997; Stufflebeam, 2003). A commonly utilised, contextual framework for programme evaluations in educational settings is Stufflebeam's Context-Input-Process-Product (CIPP) model of programme evaluation (Stufflebeam, 1967, 1969, 1971, 1973, 1983, 2003). It serves as the primary evaluation framework for this PhD thesis.

### **3.1.1 CIPP model of programme evaluation: An educational evaluation model.**

The CIPP model is an evaluation framework developed in 1967 by Stufflebeam to improve accountability in school-based programmes, but it has also been used to evaluate community-based programmes, projects, personnel, institutions, and evaluation systems (Stufflebeam, 1971, 1983). The CIPP model defines evaluation as:

the process of delineating, obtaining, providing, and applying descriptive and judgmental information about the merit and worth of some object's goals, design, implementation, and outcomes to guide improvement decisions, provide accountability reports, inform institutionalization/dissemination decisions, and improve understanding of the involved phenomena.

(Stufflebeam, 2003, p. 34)

It delineates four main purposes for conducting a CIPP evaluation (guiding decisions, providing reports, informing decisions, and promoting understanding) and four key tasks to be undertaken (delineating, obtaining, providing, and applying information). Originally, quantitative methods were employed to collect data in each of the four evaluation categories, but, over time, a mixture of methods have been used to successfully complete the context, input process, and product components required. Mertens and Wilson (2012) define these components as follows:

Input Evaluation: A needs assessment that determines what can be achieved based on the given set of goals in terms of scheduling, staffing and budget (p. 560)

Context Evaluation: Examining and describing the values, goals, missions, objectives and priorities of a program; assessing needs; and determining whether the defined objectives will be responsive to the identified needs (p. 557)

Process (Implementation) Evaluation: An evaluation that continually informs the management and main stakeholders of an ongoing intervention about early indications of progress (or lack of progress) in achieving results of a project or program, or other kind of support to an outcome (p. 561)

Product (Outcome) Evaluation: An evaluation that measures, interprets, and judges the achievement of a program in attaining its overall goals (p. 561)

After much consideration and comparison with other possible programme evaluation models (including the comprehensive, mixed-methods participatory evaluation model (Nastasi & Hitchcock, 2009), the utilisation-focussed programme evaluation (Patton, 1997), and the practical participatory evaluation model (Cousins & Whitmore, 1998)), the CIPP approach was selected as the main evaluation framework for this study because it was comprehensive, flexible, and easy to interpret, and because it also included the four components of evaluation theory and practice – input, context, process and outcome – that were of interest to the programme delivery and funding team. Also the CIPP approach was appropriate for use with a mixture of survey style data collection methods, and had been designed specifically for conducting programme evaluations (Stufflebeam, 1967). It served as the main framework for the design of both the Kiwi Can implementation and outcome evaluation studies.

The design of the implementation, and to a lesser extent, the outcome evaluation components of this study were also informed by the work of Durlak and DuPre (2008), who in the early 21<sup>st</sup> century, explored the relationship between implementation and outcomes in over 500 programme evaluation studies, and defined the notion of an implementation evaluation as “what a program consists of when it is delivered in a particular setting” (p. 329). Dane and Schneider (1998) argue that implementation evaluation involves five variables – fidelity, dosage, quality, participant responsiveness, and programme differentiation, and these were later defined by Durlak and DuPre (2008). Fidelity, they argue refers to how closely the

programme delivery adheres to the programme curriculum, dosage refers to the quantity and strength of the programme, quality refers to how well the components are delivered, and participant responsiveness and programme differentiation refer to how engaging and unique the programmes theory and practices are. They then used the variables to conduct a meta-analysis of 59 programme evaluation studies in the United States, reporting positive and significant relationships between the level of implementation across all five variables and outcomes for participants in 76% of the studies reviewed (Durlak & DuPre, 2008). This led Cross, Gottfredson, Wilson, Rorie, and Connell (2010) to argue that the implementation and outcome of programmes are inextricably linked, but without clear evidence that the programme is being delivered as intended, it is futile to measure outcomes and associate them with involvement in the programme. Gottfredson and Gottfredson (2002) argue that this is particularly true for programmes like Kiwi Can, which operate under real world conditions. Two evaluation studies were therefore designed for Kiwi Can – an implementation, and an outcome evaluation. These studies, and the investigation of health education curriculum delivery in New Zealand primary and intermediate schools embedded in this research, employed similar survey-style methodology. The theory and rationale for these choices are outlined below.

## **3.2 Experimental Design and Sampling**

### **3.2.1 Study design.**

This thesis employed two different survey designs. One was cross-sectional and non-experimental while the other was a longitudinal, quasi-experimental design. Both of these designs have different strengths and weaknesses.

Cross-sectional survey studies involve the collection of quantitative and/or qualitative data from a sample of participants at a single point in time, and are generally quick and easy to administer. They provide a snapshot of what is occurring within that sample, but only at the time the data were collected (Creswell, 2008). Researchers select cross-sectional study designs if they have limited time and resources, if they want to minimise threats to internal validity, if they want to gather a large amount of data about current attitudes and practices, or analyse and compare it using a variety of inferential statistics. Researchers who wish to investigate individual change, track

change in social processes, or investigate the direction or magnitude of a causal effect employ longitudinal studies. Longitudinal studies involve multiple data collection points. Researchers collect data from one or more samples over an extended period of time, and data collection periods can vary from days to years (Creswell, 2008). Longitudinal studies are often expensive, logistically challenging, and time consuming to administer, and are vulnerable to threats to internal validity like sample mortality (participants failing to provide data) and participant attrition (Cohen, Manion, & Morrison, 2000).

Cause and effect can be explored through the use of experimental research designs. The two most common are randomised control trial (true experimental) or quasi-experimental designs. A randomised control trial involves the random assignment of individuals or groups to an experimental condition, and is the apogee of experimental research design (Campbell & Stanley, 1963). Randomised control trials enable researchers to actively control for potential sources of bias, increasing the power and validity of results (Creswell, 2008). Where random participant allocation is not possible, researchers opt for a quasi-experimental study design, which despite the absence of a full randomisation procedure, does include a treatment and non-treatment group, pre-treatment and post-treatment measures, and provides an explicit model of the expected difference in outcomes between groups (Kenny, 1975b). Studies that employ a quasi-experimental design are however vulnerable to bias, and threats to internal and external validity (e.g., passage of time issues, participant maturation, regression towards the mean, participant mortality, participant characteristics, etc.) and are therefore only used when full randomisation is not possible (Creswell, 2008). This occurred in the outcome evaluation study of Kiwi Can presented in this thesis.

Alternatively, cross-sectional survey designs were selected for the implementation evaluation studies presented in this research as the intent was to elicit data about current practice in health education delivery and in the delivery of the Kiwi Can programme, and there was no intent or need to track changes over time. In these two studies the surveys were anonymous and designed to protect the identity (anonymity and confidentiality) of the study participants. A quasi-experimental design was selected for the outcome evaluation study presented in this research as it was not possible to randomly allocate participants to experimental conditions (participants

were either children in Kiwi Can programme schools or children in schools without the Kiwi Can programme, and due to zoning restrictions etc., it was not possible or appropriate to move them). Anonymity in the outcome evaluation study was not possible as participants were asked to respond to the same survey at two points in time and surveys had to be matched to effect comparisons. All participants formally consented to the processes involved (see Appendix B for a copy of a sample consent form).

The measurement tools (surveys) in all three studies were very similar, irrespective of study design. All three were composed of a series of close-ended questions, though the implementation evaluation and health education delivery surveys also featured open-ended comment boxes (included to encourage participants to elaborate on their responses and provide additional contextual and/or factual information). Quantitative survey-style questions were used due to concerns about time, and the ability of participants to respond to open-ended questions effectively. In all the surveys, the close-ended questions were posed using traditional, Likert-style response scales. These varied in structure from balanced five point scales to positively packed, six-point scales.

### **3.2.2 Response scale design and analysis.**

The issue of how to structure participant responses is complex (Gable & Wolf, 1993). There are a number of formats available for structuring response options and they have different psychometric properties (Brown, 2004). At present, there is little agreement about the optimal number of response categories in rating scales (Preston & Colman, 2000), though Comrey (1978) and McKelvie (1978) argue that item-response scales with five to seven response points are optimal, as scales of this length demonstrate superior reliability, discriminative power, and validity compared to scales with less than four or more than seven items. Other researchers corroborate these findings, reporting maximum reliability in scales with seven points (Finn, 1972; Ramsay, 1973), five points (Jenkins & Taber, 1977; Lissitz & Green, 1975), and both four and six points (Chang, 1994). Researchers also report that item-response scales demonstrate superior psychometric properties if they are designed without a neutral midpoint (Doyle, 1975; Ory & Wise, 1981), when the end points (Dixon, Bobo, & Stevick, 1984) and intermediate response options (Lam & Klockars, 1982) are clearly

labeled, and when verbal anchors (adjectives) that represent ascending values on the response-format continuum are included (Brown, 2004). In a balanced response scale, these adjectives are often mirrors of one another, but in positively and/or negatively packed scales, they are not (Brown, 2004).

### 3.2.3 Scales.

Positively packed scales have a larger number of response options representing the positive end of the response continuum. They are used when researchers expect respondents to be positively oriented toward the construct or idea being rated (Brown, 2004), and because they enable greater discrimination between the observations located at that end of the continuum (Hancock & Klockars, 1991). Negatively packed scales are used for the same (but opposite) reasons. Balanced scales have an equal number of positive and negative response options, and are used to reduce the tendency of respondents to agree with items regardless of content (Kline, 1993). Both packed and balanced scales produce adequate variance and have suitable psychometric properties (Brown, 2004; Deneen, Brown, Shroff, & Bond, 2013; Lai, 1994), and both can be structured to elicit information about the self-reported frequency of an attitude, behaviour, or event (called a *frequency scale*), or the degree to which an individual agrees with a statement (called an *agreement scale*).

The relative merits of frequency and agreement scales are well documented in the literature. Frequency scale formats are often criticised because of their dependence on accurate participant recall, and the threat to validity that results. Schacter (1999) argues that it is unrealistic to expect participants to accurately recall behavior frequency given the susceptibility of human beings to memory weakness, noting that transience, inattention, blocking, misattribution, suggestibility, bias, and persistence can all interfere with the accuracy of responses. He also suggests that participants may lack the episodic memory required to accurately report on a specific behaviour. Brown (2004) argues that agreement scale formats are less prone to memory error because they elicit information based on present rather than past knowledge, beliefs, opinions, or attitudes.

In this PhD, the two cross-sectional survey studies used conventional, balanced five-point Likert scales. The Kiwi Can implementation study used agreement-style rating

scales only, whereas the evaluation of health education delivery study used both agreement and frequency scale formats (though given concerns about possible bias and reporting errors, the number of frequency scales was kept to a minimum). The longitudinal, quasi-experimental design selected for the outcome evaluation study used a positively-packed, six-point scale format in anticipation of participants' positive orientation and attitude to the programme, and the desire to differentiate, as much as possible, the degree to which they expressed agreement with each statement. In all three surveys neutral midpoints were avoided (Doyle, 1975; Ory & Wise, 1981), and end points and intermediate response options were clearly labelled as recommended by Lam and Klockars (1982).

### **3.2.4 Sampling.**

Sampling is the process of selecting a sub-group of study participants from a larger population. Samples are used to facilitate data collection because it is not often possible to include an entire population of people in a single study, and Cohen, Manion, and Morrison (2000) argue that samples must be selected carefully to ensure that they embody the characteristics and natural variation present within the original population. In survey-style research, the sample is typically drawn from a large population, and is done so in accordance with a sampling frame. The frame is designed by the researchers after careful consideration of three sampling characteristics – sample size, representativeness, and access (Cohen, Manion, & Morrison, 2000).

While there is no single method for determining sample size in survey studies, Borg and Gall (1979) suggest that a valid study should number at least 100 in each major subgroup, and 20-50 in each minor subgroup. They also suggest that the level of accuracy (sampling error) and level of probability be considered when establishing sample size. Sampling error is a measure of the natural variation in responses that arise between the sample and the total population. It is often interpreted as the difference between the population mean and the sample mean (Cohen, Manion, & Morrison, 2000). The size of the sampling error typically decreases as the sample size increases (Hopkins, Hopkins, & Glass, 1996). A number of calculators now exist that compute sample size based on a measure of the size of the total population, the sampling error, and the confidence interval associated with the sample. For this

research, a calculator produced by Raosoft (<http://www.raosoft.com/samplesize.html>) was used to help determine the sample sizes for all three studies, though the final sizes were larger than than the calculator suggested, as the researcher sought to mitigate the potential for attrition and respondent mortality (Blalock, 1979).

In order to draw valid conclusions from the results of survey studies, study participants must be matched as closely as possible to the entire population on all demographic characteristics of import (Cohen, Manion, & Morrison, 2000). If there are significant differences, than the size of the sampling error will increase and diminish the power the statistical tests used and the veracity of conclusions drawn (Evans, 1992). In this PhD, the sample participants were all representative of the original population from which they were drawn.

Finally, access to participants is also an important consideration for researchers. Cohen, Manion, and Morrison (2000) argue that sample access can be problematic when there are financial, logistical, or personal constraints present, and limited or difficult to secure when children are involved. In this PhD, access to children participating was secured with the co-operation of schools and families.

#### ***3.2.4.1 Sampling framework.***

Once the issues related to access, representation, and size have been considered and resolved, researchers must establish the overall design of the sampling framework. There are two main methods of sampling – probability (or random) sampling, and non-probability (or purposive) sampling (Cohen & Holliday, 1979, 1982, 1996). In a probability framework, everyone in the population has an equal chance of being randomly selected, whereas in a non-probability framework, constraints are imposed to purposely include or exclude selected members of the population (Cohen, Manion, & Morrison, 2000). Probability samples are naturally less susceptible to bias, and are used when researchers wish to generalise results to the wider population. Non-probability samples are used when researchers are unable to use a random sampling technique, or when they do not wish to generalise. The most common types of probability samples are – random samples, systematic samples, stratified samples, cluster samples, stage samples, and multi-phase samples. The most common types of

non-probability samples are convenience samples, quota sample, purposive samples, dimensional samples, and snowball samples (Cohen, Manion, & Morrison, 2000).

In this PhD, two different techniques were used to establish the samples. A random sampling technique was used in the investigation of health education delivery in primary and intermediate schools, and all the primary and intermediate school health education teachers in New Zealand were invited to participate. In the Kiwi Can evaluation studies, participants were selected using a purposive, stratified, cluster sampling technique. Each of the eight Kiwi Can regions were assigned a stratum and data were collected from all eight participating Kiwi Can regions (stratums) around New Zealand, though from purposively selected schools in each stratum rather than individual participants – a decision made because of time and financial constraints, and with an awareness that the homogeneity of responses from one region may skew results slightly (Sapsford, 2007). Once the samples had been identified, schools and children were invited to participate. Those who agreed to participate in the health education curriculum study were sent electronic surveys while those who agreed to participate in the Kiwi Can implementation and outcome evaluation studies were sent hard surveys. Completed surveys were returned to the University of Auckland for data preparation and analysis.

### **3.3 Data Preparation**

A number of studies have reported issues with response validity on self-report surveys with adolescents (Fan et al., 2006; Sudman & Bradburn, 1974, 1982), so in this research all of the survey data were screened and cleaned using a series of identical preparatory procedures before analysis. In an investigation of adolescent self-reported health status in the United States, Fan et al. (2006) identified carelessness, confusion, lack of effort, and intentional mischief as reasons for inaccurate responses. Children who had provided responses indicative of a lack of effort or intentional mischief were termed *jokesters* by the research team, while those whose responses appeared careless or confused were termed *inaccurate responders*. While the distorting effect of the inaccurate responders was of only minor concern, the responses provided by the jokesters introduced significant systematic error into the study and threatened the internal validity of the results. As a consequence, the decision was made to manually screen the data in this study for evidence of blatantly mischievous or misleading

reporting (e.g., surveys which featured mono-numeric responses or where respondents drew visible patterns on the pages) and these surveys were withdrawn. The remaining surveys were then screened for evidence of missing data using two techniques – a manual inspection of the data set, and a systematic evaluation of the distribution of missing values. Although there are no firm guidelines as to the amount of missing data that can be tolerated for a given sample size (Tabachnick & Fidell, 2007), in this PhD cases with > 10% missing responses were deemed to have too much missing data to warrant imputation and were removed. The remaining data set was then evaluated to determine the nature of the missing data distribution for cases with  $\leq 9.99\%$  missing data.

Data can be Missing Completely at Random (MCAR), Missing at Random (MAR), or Missing Not at Random (MNAR; Tabachnick & Fidell, 2007). In this PhD, Little's MCAR test can be used to determine whether the data are missing completely at random (Tabachnick & Fidell, 2007). A statistically non-significant result ( $p \geq .05$ ) provides support for the hypothesis that the pattern of missing values does not diverge from randomness. An ANOVA using Hochberg's GT2 post hoc test can also be used to determine if data are missing at random when group sizes differ significantly (Field, 2013a), and was used in this thesis to determine whether there were statistically significant differences in the mean responses for surveys with no missing data (0%) versus those with missing data (1-10%). A statistically non-significant result ( $p \geq .05$ ) indicated that the data were missing at random (Field, 2013a). Non-significant results from both tests provided support for the use of an imputation procedure to create a complete data set. In each of the three studies in this PhD thesis, the results of the Little's MCAR and ANOVA tests indicated that the data were missing completely at random.

### **3.3.1 Imputation.**

While there are a number of imputation procedures available (Acock, 2005), all missing data in this thesis was imputed using the expectation maximization (EM) procedure in the SPSS 21 software programme. EM is an iterative process during which values are imputed successively until the current underlying covariance matrix for the iteration is almost identical to the covariance matrix from the previous iteration (Acock, 2005). In this PhD, the single imputation EM method was selected because

there was very little missing data, the data were missing completely at random, there was assumed likelihood under the normal distribution, and the data were normally distributed according to the skewness and kurtosis values (Tabachnick & Fidell, 2007). Skewness and kurtosis are measures of normality for both univariate and multivariate data. Univariate data are data with a single outcome variable, while multivariate data involves the measurement of two or more outcome variables (Field, 2013a). Univariate normality is characterised by a distribution with a single peak or mode, a large proportion of cases close to the mean (i.e., the classic normal distribution has cases spread around the mean so that there the peak of the distribution ensures only 68% of cases are within +/-1 standard deviation of the mean), and cases being centred in the middle of the response range. Univariate normality is a necessary, though not sufficient, condition for multivariate normality. Multivariate normality is an extension of univariate normality and is a probability distribution for a set of variables (Field, 2013a).

Field (2013a) argues that skewness is a general measure of the symmetry of a frequency distribution. A skew score of 0 indicates that the distribution is perfectly symmetrical. Positive skew values indicate that majority of scores are distributed below the mean, while negative skew scores indicate the opposite. Kurtosis, on the other hand, is a measure of the degree to which scores for a particular item cluster in the tails of the frequency distribution. Positive kurtosis values ( $> 0$ ) indicate that a distribution is too peaked, while a negative kurtosis ( $< 0$ ) score indicates that a distribution is flat. Data are assumed to be normally distributed if the standardised skewness values falls between 0 and +/- 2.0, and standardised kurtosis values ( $B_2$ ) do not exceed +/-7.0 (DeCarlo, 1997; West, Finch, & Curran, 2005). Critical ratios of the statistic relative to its standard error are expressed as z-scores (i.e.,  $M = 0$ ,  $SD = 1$ ) and values  $> 5.0$  indicate non-normality that may require intervention (Bentler, 2005).

Multivariate statistical techniques (e.g., factor analysis, multiple regression analyses, structural equation modeling) require that the variables are multivariate normal and have homogeneity of variance (Field, 2013a; Tabachnick & Fidell, 2007).

Homogeneity of variance refers to the degree to which the variance associated with a specific parameter (dependent variable) is similar between groups of interest in a study (Field, 2013a). It is important to assess the degree to which the variances are

homogeneous, as unequal variances can create bias and inconsistency in the estimates of standard error that are associated with parameter estimates in a model (Hayes & Cai, 2007). This, in turn, will lead to bias in confidence intervals and tests of significance. In SPSS, the homogeneity of the variance structures underlying data can be evaluated using Levene's test, with a statistically significant result ( $p \geq .05$ ) indicating heterogeneity. The Levene's test statistic is robust even when the sample size is relatively small ( $n \leq 50$ ), however, when the sample size is large the test statistic must be interpreted with caution (Field, 2013a).

If the data does not demonstrate normality and/or homogeneity of variance, Field (2013a) identifies four methods that can be used to redress this problem and reduce potential bias – trimming the data, winsorising, using robust methods, and transforming the data. He also advocates the use of a robust method that is unaffected by violations of assumptions or outliers. One such method is *bootstrapping* (Efron & Tibshirani, 1993). In bootstrapping the sample data are treated as the population from which a series of smaller samples are drawn. Means are then calculated for each bootstrapped sample along with confidence intervals and significance ( $p$ ) values for these parameters.

It was not necessary to use any of the methods listed above in this PhD thesis as the data were normally distributed and demonstrated homogeneity of variance. Therefore, the data underwent the imputation or EM process. The EM method consisted of two steps – the E and the M steps. During the first step the programme computed the conditional expectation of the missing data, given the observed values and current estimates of the parameters. The expectations were then substituted for the missing data. In the second step, maximum likelihood estimates of the parameters were computed. During this stage they were computed as though the missing data had been filled in.

Once imputed values have been accepted, the data set has to be inspected to ensure that all imputations fall within the logically expected range of responses. This is essential in rating scales which have minimum and maximum values that cannot be logically violated, and because the EM procedure may impute a plausible response below the minimum or above the maximum. Such values below need to be adjusted to the response scale minimum and maximum values. All remaining values do not

require adjustment because, as a continuous response scale, it is expected that these values are valid and accurate representations of the range of participant response.

At the end of the data screening process for each study in this thesis, there was a complete set of ordinal data ready for analysis, and the decision was made to treat this ordinal survey data as continuous. This decision was largely based on the work of Bollen (1989) and Finney and DiStefano (2006) who found that as the number of ordered categories increases, data more closely approximates continuous-level data. Further, Babakus, Ferguson, and Jöreskog (1987) have also reported that, in the world of model testing, fit indices perform well if approximately normally distributed five-category ordinal data are treated as continuous. On this basis, it was determined that the quantitative data from all survey studies in this PhD could be treated as continuous and that means could be calculated and compared both between and within groups.

### **3.4 Data Analysis**

#### **3.4.1 Model design.**

The data from the implementation and outcome evaluation of Kiwi Can was analysed using both a SEM and a CFA technique. SEM is a statistical methodology or procedure that adopts a confirmatory approach to the analysis of the structural theory underlying a data set (Byrne, 2010). It is based on two main assumptions – that the causal processes underlying a data set can be represented by a series of structural (i.e., regression) equations, and that it is appropriate and valid to represent these relationships in pictorial form (Byrne, 2010). The final product of the structural equation process is a full latent variable model that depicts the relationship between observed (i.e., manifest) and unobserved (i.e., latent) variables. It also includes measures of the error (i.e., residuals) associated with the relationships. The full latent variable model is composed of both a measurement model and a structural model. The measurement model depicts the relationship between latent variables and their observed variables. The structural model depicts only the relationships that exist between the latent variables themselves. The measurement model is often called a CFA model because it is used to confirm that a specific factor structure underlies the data. More specifically, it is used to confirm an existing theory that posits that an underlying latent variable, or factor, is strongly associated with one or more observed

variables. The relationship between latent and manifest variables can also be assessed using an EFA approach. Unlike the CFA, an EFA is not used to test theory. Instead, it is used when there is little or no theoretical understanding about the relationship that exists between latent and manifest variables.

In this PhD study, the predicted factor structure underlying the Kiwi Can implementation and outcome data were investigated using a CFA approach. CFA was chosen over EFA because the researcher had a theoretically driven hypothesis about the underlying structure of the data (Byrne, 2010). Confirmation of the factorial validity of the hypothesised models was conducted using a SEM technique in the AMOS (Version 21) software programme (IBM, 2011). SEM was chosen because it was capable of testing the overall fit of complex models (and there were a number in this thesis). Each of these complex measurement and structural equation models were designed in AMOS and then tested for accuracy or *goodness-of-fit* using participant data.

#### ***3.4.1.1 Assessing model fit.***

There is debate as to how to evaluate the quality of an admissible model. Originally, model fit was assessed using an asymptomatic chi-square ( $\chi^2$ ) test statistic (Hu & Bentler, 1999). A large  $\chi^2$  relative to its degrees of freedom was evidence of a poor fitting model. Over time, concerns have been raised about the adequacy of this method, as it is vulnerable to sample size and violations of normality (Bentler & Dudgeon, 1996; Chou, Bentler, & Satorra, 1991; Curran, West, & Finch, 1996), and as a result, alternative measures of model fit termed *fit indices* have been developed. Typically, these are classified as *absolute* or *incremental/comparative* fit indices (Bollen, 1989; Gerbing & Anderson, 1993). Absolute fit indices measure how well an a priori model reproduces sample data. Examples of absolute fit indices are the goodness-of-fit and adjusted goodness-of-fit indices (Bentler, 1990), cross-validation index (Browne & Cudeck, 1989), gamma hat (Steiger, 1989), standardised root-mean-square residual, (SRMR, Jöreskog & Sorbom, 1981), and the root mean square error of approximation (RMSEA, Steiger & Lind, 1980). The incremental or comparative fit indices can also assess model fit by comparing a target model with a restricted, baseline model. The baseline model typically involves the design of a model with uncorrelated observed variables (Bentler & Bonett, 1980). Examples of common

incremental fit indices include the normed fit index, NFI (Bentler & Bonett, 1980), the non-normed fit, NNFI or TLI (Bentler & Bonett, 1980), and Bentler's comparative fit index (CFI, Bentler 1989).

Both absolute and incremental fit indices were designed to avoid problems with sample size and distributional mis-specification, but they have not been without problem. Hu and Bentler (1999) argue that fit indices can produce conflicting results about the extent to which a model fits the observed data. Applied researchers must therefore make informed decisions about the selection and the use of fit indices in their work. In the literature, the most robust fit indices (i.e., those that are immune to sample size, model complexity, and even model misspecification) are – the  $\chi^2$  per  $df$  ratio (Marsh, Hau, & Wen, 2004), SRMR (Bentler, 1995; Fan & Sivo, 2007), gamma hat (Fan & Sivo, 2007), RMSEA (Steiger, 1990), and the CFI (Bentler, 1990). It must be noted that the RMSEA is sensitive to model complexity and will become inflated when used with highly complex models (Breivik & Olsson, 2001; Fan & Sivo, 2007). Further, the CFI tends to reward simple ( $\leq 3$  factors) models (Fan & Sivo, 2007). Therefore, in this PhD thesis, the fit of the measurement and structural equation models was assessed using the  $\chi^2/df$ , CFI, RMSEA, SRMR, and gamma hat indices (Hu & Bentler, 1998, 1999; Marsh, Hau & Wen, 2004). A model was deemed to have 'good' fit if the fit values were:  $< 3.00$  ( $\chi^2/df$ ),  $> .95$  (CFI),  $< .05$  (RMSEA),  $< .06$  (SRMR) and  $> .95$  (gamma hat). Alternatively, models were deemed to have 'acceptable' fit if the values were  $< 3.83$  ( $\chi^2/df$ ),  $> .90$  (CFI, gamma hat), and  $< .08$  (RMSEA, SRMR; Hu & Bentler, 1998, 1999; Marsh et al., 2004.).

#### ***3.4.1.2 Model misspecification.***

At times, SEMs may not be accurately specified. Model misspecification can be detected in the standardised residuals modification indices (MIs) and expected parameter change (EPC) output in AMOS. According to Jöreskog and Sorbom (1993), standardised residual values of  $> 2.58$  indicate that there is a statistically significant misfit between two variables in the model. Large MI and EPC values also indicate a potential misfit between variables. Byrne (2010) recommends that both the residual and MI/EPC data be carefully considered and that parameters be removed from the model if they have large MI and EPC values, and are not substantively meaningful. In this PhD thesis, the standardised residuals, MI and EPC values were

considered, and in some cases parameters were removed in order to reduce model misspecification.

### ***3.4.1.3 Model selection.***

Multiple models can fit equally well to the same data. This means that a combination of statistical and theoretical approaches is needed to select the most defensible model, and need to be robust enough to ensure that the final model is the most accurate representation of the factor structure underlying the observed data – otherwise, the results generated in all subsequent data analysis processes will be invalid. To determine whether one model fits better than another, the difference in  $\chi^2$  as a ratio of the difference in  $df$  can be examined. Since  $\chi^2$  is a normally distributed statistic, it can be converted into a z-score, which permits determination as to whether the difference is statistically significant (Wilson & Hilferty, 1931). A statistically significant difference indicates that the reference model is a better fit to the data. Inspection of the difference in the Akaike Information Criterion (AIC) can also assist in identifying models which are closer to the data. Models with difference in AIC ( $\Delta_i$ ) values of  $< 2$  have substantial support,  $4 < \Delta_i < 7$  have less support and  $\Delta_i > 10$  have no support (Burnham & Anderson, 2004). In this PhD thesis, both the difference in  $\chi^2$  and AIC ( $\Delta_i$ ) values were used to select the appropriate structural model.

### ***3.4.1.4 Invariance testing.***

Once an adequately fitting model has been established, it should be tested for statistical equivalence (i.e., invariance) across participants who vary according to demographic or other relevant attributes. Invariance testing evaluates whether the model operates similarly, or can be replicated, for all groups of interest in a study (Byrne, 2010). It is necessary to establish invariance in a model before it can be used as the basis for exploring group differences in mean response (Byrne, 2010).

Essentially, invariance testing establishes the degree to which a model has equivalent parameter characteristics between groups (Cheung & Rensvold, 2002). In order to establish invariance, the configuration of the paths within a model must be similar. In a multigroup situation, this is accepted if the RMSEA is  $< .05$  (Cheung & Rensvold, 2002). Most invariance testing begins with a global test of the equality of covariance

structures across all groups of interest (Jöreskog, 1971), termed *unconstrained invariance testing* (Byrne, 2010). Models that demonstrate unconstrained invariance are then subjected to increasingly stringent and restrictive tests of invariance. In each level, specific parameters are constrained and model fit is assessed. In most cases, the factor loadings in a measurement model are constrained (i.e., they are made equal) first. This level of invariance testing is termed *measurement invariance* or *measurement equivalence*, and followed by constraints of the factor co-variances (*scalar invariance*), and the structural regression paths (*structural invariance*) (Byrne, 2010). In some cases, the latent mean structures present in the data are also assessed for invariance (Byrne, 2010).

A model is considered to be equivalent if the fixing of a parameter category to equal values for all groups creates a change ( $\Delta$ ) in the CFI of  $< .01$ , or if the  $\chi^2$  difference test for the constrained model is not significant ( $p \geq .05$ ). Sequential testing of equivalence stops if the difference in CFI is  $> .01$  or the  $\chi^2$  value is significant ( $p \leq .05$ ).

In this PhD, the  $\Delta$ CFI criterion was selected over the traditional  $\chi^2$  difference test because the traditional approach has proven to be excessively stringent under structural modeling conditions (Cudeck & Browne, 1983; MacCallum, Roznowski, & Necowitz, 1992). Further, Cheung and Resvold (2002) argue that it can be used to test the invariance of models that have been designed and tested using cross-sectional and/or longitudinal, repeated measures data. In a repeated measures study the data are collected from the same group of individuals, using the same instrument, at multiple time points.

#### ***3.4.1.5 Modeling change over time.***

In SEM it is now conventional to analyse repeated measures data using a latent growth curve model approach that estimates the time-varying effect of an initial intercept (i.e., starting score) parameter, and a slope (i.e., rate of change) parameter, on the repeated scores (Curran & Bollen, 2011). Given that this involves the calculation of the underlying mean and covariance structures for the intercept and slope, this approach may not be estimable with fewer than three data collection phases (Curran & Bollen, 2011). In contrast, autoregressive or Markov simplex modelling,

based on simple regression principles, can be used to estimate change when only two data points are available. Markov simplex models suggest that performance at any one time is determined by performance at the time immediately preceding – that scores at Time 2 are strongly influenced by performance at Time 1. When two different, but correlated, scores are tracked over time, it is possible that each score influences not only itself, but also the other construct at the next instance of data collection. This approach is called *cross-lagging* (i.e., variables predict themselves and the other construct at the next time period). Thus, in a bivariate cross-lagged autoregressive model, each construct is predicted by the immediately preceding instance of itself, and the other construct.

In this PhD, autoregressive, cross-lagged (AR-CL) panel models were used (Finkel, 1995; Hertzog & Nesselroade, 2003; Jöreskog, 1979). The autoregressive or univariate simplex model illustrated the relationship between the same latent constructs, regressed onto one another, while the cross-lagged portion provided information about the structural relationship between different latent constructs over time (Finkel, 1995; Hertzog & Nesselroade, 2003; Kenny, 1975a). The stability coefficients (correlations) produced in the cross-lagged portion of the model indicated the degree to which variation in one latent construct was predictive of change in a second latent construct over time (Kenny & Harackiewicz, 1979). AR-CL models are useful for modelling change at the group level only, do not account for absolute change in individual scores, and are unable to compute rates of individual change over time (Hertzog & Nesselroade, 2003; Kline, 2005). Despite these potential weaknesses, the AR-CL approach was selected for this study because it was not appropriate to use more elegant or complex latent growth curve modelling with only two occasions of measurement (Hertzog & Nesselroade, 2003; Rogosa, Brandt, & Zimowski, 1982). This is because the growth curve models generate individual, continuous growth trajectories based on repeated measures data. It is most accurately computed using three or more data points (Curran & Bollen, 2011).

Once the structural equation models have been designed, tested, and demonstrated that they have adequate fit and invariance, differences in participant responses, are compared using a variety of statistical techniques. In this PhD, four specific techniques were employed, and are presented in the following order: (i) cluster

analysis; (ii) paired samples t-tests; (iii) analysis of variance (ANOVA), factorial analysis of variance, multiple analysis of variance (MANOVA), and repeated measures analysis of variance (RM-ANOVA); and (iv) multiple regression analysis. With the exception of the cluster analysis, the techniques were used to identify statistically significant differences in mean responses for participants in each of the survey studies. In each analysis, effect sizes were also computed to establish the magnitude of the mean difference. In this PhD, Cohen's  $d$ , Adjusted  $R^2$  and  $f^2$  values were used to ascertain effect size differences. The theoretical basis for these effect size measures are described at the end of this section.

### **3.4.2 Statistical data analysis procedures.**

#### ***3.4.2.1 Cluster analysis.***

In Part B of the implementation evaluation survey for Kiwi Can, participants were asked to respond to a series of questions using a binary or dichotomous response format (yes/no). While this is a common format in survey style research, it is not without fault. Burns (2000) argues that the dichotomous response format can be problematic and threaten the construct validity of the analysis because it fails to capture the true variability and scope of participant experience. Under a yes/no format, individuals are forced to select the option that is closest to the 'truth', but may not accurately reflect their lived experience. Further, binary question formatting produces a final data set that has a bimodal frequency distribution which violates the assumption of normality (Field, 2013a). This renders the use of standard parametric tests, based on the normality assumption, inappropriate. As a result, non-parametric approaches must be taken.

In this PhD, the cluster analysis technique was selected for use under these conditions. It was used in Part B of the implementation evaluation of the Kiwi Can programme (see Chapter 5) to test the hypothesis that responses about the frequency of recall regarding i) learning and ii) skill development in the Kiwi Can programme were indeed distinct categories. It was necessary to test the hypothesis that there were two distinct categories of responding before any further analyses were undertaken. Cluster analysis is a generic term used to describe a variety of classification procedures. It became a popular method for classifying items in the biological sciences in the 1960s

(Sokal & Sneath, 1963). Over the years it has been adopted for use by many in the social science field, including education (Aldenderfer & Blashfield, 1984).

In order to test this hypothesis, half of the original data set ( $n = 160$  cases) was randomly selected and subjected to a k-means cluster analysis (iterative partitioning) test. In this analysis, the k-means or k-pass cluster technique was chosen over traditional hierarchical methods because there was a large number of cases ( $n = 319$ ) and, unlike the hierarchical method, the k-means approach does not require the calculation and storage of a similarity matrix (Aldenderfer & Blashfield, 1984). Generally, it is also more stable and capable of overcoming problems with poor initial partitions (Gower, 1967). In fact, the k-means approach can modify a poor early partition while the single-pass hierarchical method cannot. It must be noted that the k-means analysis can be susceptible to a “problem of local optima” (Aldenderfer & Blashfield, 1984, p. 48), though Milligan (1980) has proven that this can be negated by using an initial starting partition from average linking clustering. To overcome these issues, the average linking cluster approach was used to generate the original partition. The product of the k-means cluster analysis is termed a cluster solution, and includes the number of clusters that arose from the analysis and the membership of each binary item to a specific cluster. Scores of ‘1’ indicate membership in a cluster, while scores of ‘0’ indicate non-membership. A measure of the distance between the centroids of each cluster is also provided. In this PhD, the validity of the original cluster solution was assessed using the replication procedure. The second half of the original data set ( $n = 160$ ) was subjected to an identical k-means cluster analysis, and the results of the two analyses were compared. The results were identical. This internal validation process was first introduced by Goldstein and Linden in 1969 (Aldenderfer & Blashfield, 1984), and was selected for this study because it is considered superior to the traditional cophenetic correlation procedure, and to the use of significance tests. In fact, Aldenderfer and Balshfield (1984) strongly caution against testing for the significance of clusters using a multivariate analysis of variance (MANOVA) or multiple analysis of variance procedures. This is because cluster analysis separates entities into distinct clusters that have little to no overlap, and conducting significance tests on distinct entities will naturally produce a highly positive and statistically significant result. This result can appear in the absence of a

statistically significant difference (in clusters), and can lead researchers to make a Type I error.

The cluster analysis technique was the only non-parametric test selected for use in this PhD. The remaining tests were all parametric, and selected over non-parametric options because they are more powerful, and therefore more capable of identifying statistically significant differences between variables (Burns, 2000). The purpose, advantages, and rationale for the selection of these tests can be found below.

### ***3.4.2.2. Paired samples t-tests.***

A t-test is used to test for statistically significant differences in observed sample means under two experimental conditions (Field, 2013a). It is used to test the null hypothesis that there is no difference between the population groups and that any observed variation is due to chance. The statistical t-test procedure generates a test statistic and accompanying *p*-value that indicate the likelihood that the results were not obtained by chance (*p*-value is  $\leq .05$ ). Under these conditions the null hypothesis is rejected. All t-tests are based on a number of assumptions, including a normally distributed population, a dependent variable measured on a continuous, interval or ratio level, random sampling of the data, and homogeneity of variance (Field, 2013a). In a t-test, the independent variable must be categorical.

The data for a t-test can be collected from different participants or from the same participants. If the data is collected from different participants, an independent samples t-test is used. If it is collected from the same participants, a paired samples t-test is used (Field, 2013a). In this PhD study, the data were always collected from the same participants, and therefore paired samples t-tests were used when there were only two mean scores to compare between individuals in the same group. The t-test is the most appropriate tool for comparing mean scores because it is able to produce estimates of the direction and magnitude of all statistically significant differences. It is not, however, able to compute mean differences between three or more groups. In such cases, a different technique has to be used. In this PhD, the one-way analysis of variance technique (ANOVA) was used. Variations of the ANOVA (i.e., multivariate ANOVA, repeated measures ANOVA, and factorial ANOVA) were also employed under different conditions.

### ***3.4.2.3 Analysis of variance testing – ANOVAs.***

The analysis of variance technique (ANOVA) is a collection of different statistical models that can be used to determine whether significant differences exist between means from three or more groups. One-way ANOVAs are used to explore differences in a single dependent variable between multiple groups. ANOVAs are used in lieu of multiple t-tests because they reduce the likelihood of making a Type I error. Type I errors occur when a false null hypothesis is rejected (Evans, 2014). One-way ANOVAs produce an omnibus test statistic ( $F$ ), and a corresponding measure of significance ( $p$ ) (Burns, 2000). The omnibus statistic is computed using a ratio of systematic to unsystematic variance in mean scores. A  $p$  value of  $\leq .05$  is commonly used as the cut-off point for determining the significance of the  $F$  statistic, however, it can be adjusted (Field, 2013a).

One-way ANOVAs are based on the assumption of normality of data and the homogeneity of variance. If these assumptions are violated, it is possible to adjust the  $F$ -test statistic to correct for the problem. The two most common corrected versions of the  $F$ -test are the Brown-Forsythe  $F$  (Brown & Forsythe, 1974), and Welch's  $F$  (Welch, 1951). In this PhD, neither statistics were used because there were no significant concerns about the homogeneity of variance or normality of the data. Instead, the standard  $F$ -test was used to indicate significant difference. The  $F$ -test does not indicate the direction of the mean difference, nor does it indicate where the difference lies. Post hoc tests must be used to generate this information. In the case of a one-way ANOVA, there are a number of possible post hoc procedures to choose from.

### ***3.4.2.4 Post hoc procedures – ANOVAs.***

Post hoc tests consist of a series of pairwise comparisons. As such, they are vulnerable to problems controlling the Type I error. Most post hoc tests are able to control the Type I error rate effectively. Some reportedly perform poorly when group sizes are unequal, and when population variances are different. According to Field (2013b), the most robust test procedures for unequal group sizes are Hochberg's GT2, Gabriel's pairwise test, and the Games-Howell procedure. Of the three, the Games-Howell is the most powerful and the most accurate when sample sizes are unequal,

and when the equality of group variances cannot be assured (Field, 2013b). It can be too liberal, however, when sample sizes are small. In this PhD, there were large but unequal sample sizes, and it was not possible to ensure equality in group variances, therefore the Games-Howell post hoc procedure was selected for use with each of the one-way ANOVAs. One-way ANOVAs test for statistically significant differences in means when there is one dependent variable. Under conditions with two or more dependent variables, a multiple analysis of variance technique must be employed. This is used in lieu of multiple, separate ANOVAs because by including all dependent variables in a single analysis, the risk of making a Type I error by taking into account the relationship between dependent variables is reduced.

#### ***3.4.2.5 Multiple analysis of variance tests.***

A multiple analysis of variance (MANOVA) procedure operates on the same principles as one-way ANOVAs. It too assumes that the data were randomly sampled, that the data are normally distributed, and that there is homogeneity of variance. However, with a MANOVA, homogeneity of variance must be assumed for each dependent variable and for the correlation between any two dependent variables in the same group. This final assumption is assessed using the population variance-covariance matrices. Covariance testing is undertaken in SPSS using Box's test of equality of covariance matrices. A non-significant test result indicates that the matrices are similar. In this PhD, the results of Box's test were significant in a few cases. It was expected that the significant result was a product of the large sample size and not a reflection of genuine invariance because Box's test has proven to be unstable with large sample sizes (Tabachnick & Fidell, 2007). However, in order to verify this assumption, the Q-Q plots and covariance structures present in the data were examined manually for evidence of invariance. Tabachnick and Fidell (2012) argue that the variance within groups and the size of the groups must be compared. If there is larger variance in the larger groups, than in the smaller groups, it is likely that the probability values will be conservative and the test statistics can be trusted (Tabachnick & Fidell, 2012). In each case within this PhD, there was larger variance within the smaller groups, indicating that the test statistic could not be trusted. An assumption of homogeneity was therefore made. This assumption influenced the selection of the test statistic for the MANOVAs in this thesis. While there are four

possible MANOVA test statistics to use, the decision was made to select the Pillai-Bartlett trace ( $V$ ), rather than the Hotelling's  $T^2$  trace, Roy's largest root, or Wilk's lambda, because it was more robust with unequal sample sizes and assumed homogeneity than the other three statistics (Stevens, 1979).

Akin to one-way ANOVAs, the  $F$ -statistic produced in the MANOVA is an omnibus statistic. Therefore, post hoc tests must be used to indicate the direction and magnitude of the statistically significant difference in mean scores. The Games-Howell procedure was used as the post hoc test statistic for all MANOVA tests in this PhD because sample sizes were unequal. MANOVAs were used to assess mean differences between independent groups during a single period of time. They were not used to assess mean differences between participants across time, when the same participants provided data at both time points. In these situations, repeated measures ANOVAs were used.

#### ***3.4.2.6 Repeated measures ANOVAs.***

The repeated measures ANOVA (RM-ANOVA) is a category of ANOVA used to compare mean differences in dependent variables when data are collected at multiple time points from the same participants (Field, 2013b). It was designed to account for the shared within-subjects variance (or lack of independence), that is characteristic of data collected in this manner (Field, 2013b). RM-ANOVAs are based on an assumption that there is a relationship between the pairs of data provided by the same participant at each data collection point. It is termed the *assumption of sphericity* (Field, 2013b), and refers to the equality of the differences between treatment levels. Sphericity is an issue when there are three or more conditions under investigation. It is assessed in SPSS using Mauchly's test, which was not used in this thesis because all repeated measures data were collected at only two points in time, and therefore there were only two conditions included.

The RM-ANOVAs can be one-way or factorial. A one-way RM-ANOVA is used when there is one independent variable under investigation and the data are collected from the same participants over time. A factorial RM-ANOVA is used when there is more than one independent variable and the data are collected from the same participants. In this PhD, one-way RM-ANOVAs were used in the outcome

evaluation of Kiwi Can, and factorial RM-ANOVAs were used in the analysis of the interaction between transience and school condition on individual social attitudes and classroom climate data. For all RM-ANOVAs, the Hochberg GT2 test was used as the post hoc test statistic because the Games-Howell post hoc test statistic is not robust enough to use under conditions with repeated measures. Further, the Hochberg GT2 has proven to be more accurate than the alternative post hoc test procedures (i.e., the Tukey or Scheffe post hoc tests) when sample sizes are different (Field, 2013b).

The final set of analyses undertaken in this PhD were regression analyses. They were used to explore the relationship between teacher beliefs and teacher practice. Multiple regression analysis was appropriate for this investigation because it was necessary to examine the impact of numerous beliefs on practice.

#### ***3.4.2.7 Multiple regression.***

Multiple regression is a type of standard regression analysis. Standard regression involves the design of a linear model,  $y_i = (b_0 + b_1X_i) + \epsilon_i$ , that is used to predict the value of a dependent, or outcome, variable ( $y_i$ ) from one or more independent variables ( $X_i$ ), an unstandardised parameter value ( $b$ ), and a measure of error ( $\epsilon$ ). The linear model is designed by a statistical software programme using available data. It represents an estimate of the true linear relationship between the outcome variable ( $Y_1$ ) and predictor variables (e.g.,  $X_1$ ). The accuracy, or goodness-of-fit, of the estimated model is assessed using a series of measures. The three most common measures are the  $R^2$ ,  $F$  and  $b$ . The  $R^2$  is a measure of the proportion of variance in the outcome variable that is shared by the predictor variable. The  $F$  statistic indicates the degree of variability in the outcome variable that is explained by the model, and the  $b$  value indicates the strength of the relationship between the predictor and outcome variables. The  $F$  statistic is associated with a significance value. If  $p \leq .05$ , then it is assumed that the predictor variable significantly predicts the outcome variable (Field, 2013b).

As was the case with ANOVAs, there are a number of assumptions upon which the linear model is designed and tested. Field (2013b) argues that these assumptions include the facts that: (i) the predictor and outcome variables are linearly related to one another and that the combined effect of the predictors is best described by

summing these values (additivity and linearity); (ii) the residual (error) terms are uncorrelated (independent errors); (iii) the variance of error terms for each level of the predictor values are constant (homoscedasticity); (iv) the residuals in the model are random, normally distributed variables with a mean of 0 (normally distributed errors); and (v) there is no perfect linear relationship between two or more of the predictor variables (multicollinearity). The collinearity between variables is assessed using a correlational analysis. Correlation is a measure of the degree to which two variables co-vary linearly. It is measured using a Pearson's correlation co-efficient ( $r$ ) that varies from -1 (perfect negative correlation) to +1 (perfect positive correlation). A correlation of 0 indicates that is no linear relationship between the variables. Variables that correlate at  $r \geq +/- .90$  are deemed to be too highly correlated (Field, 2013b).

In a simple linear regression there is one predictor variable. In a multiple regression analysis there are two or more predictor variables. These variables can be added into the linear model in different ways. The three most common are hierarchical (blockwise) entry, forced (simultaneous) entry, and stepwise entry. In a hierarchical regression, predictors are added into the model manually in the order of assumed importance. In a forced or simultaneous entry, the variables are added into the model all at once, and in a stepwise regression they are added into the model on the basis of mathematical criterion. The stepwise approach can involve a forward or backward method. The forward method involves adding predictors one at a time to the model, and assessing model fit after each addition. The backward method involves inserting all predictors into the initial model before comparing the contribution that each predictor variable makes to the model with a pre-determined removal criterion. Variables that meet the removal criterion (i.e., that do not contribute a statistically significant amount) are removed from the model. Field (2013b) argues that the stepwise method is the least appropriate, as it involves assessing the goodness of fit for a variable solely on the basis of the other variables in that model. This can lead to the false acceptance or rejection of a variable if the other variables entered into the model are ill-fitting or inappropriate. Further, the stepwise method is conducted entirely by computer. This not only removes all of the methodological power from the researcher but it also fails to acknowledge the theoretical importance or validity of the predictor variables.

In this PhD thesis, the simultaneous, multiple regression approach was adopted in the study of health education delivery in New Zealand schools. It was chosen because there were more than three predictor variables of import in the study, there was no expectation that specific variables were stronger predictors of the outcome, and there was a desire to produce replicable results. Once the multiple regression analyses, ANOVAs, and paired samples t-tests had been conducted, it was necessary to compute the size of all statistically significant effects.

#### ***3.4.2.8 Calculating effect size.***

Effect size is a standardised measure of the magnitude or size of the observed effect that an independent variable has had on a dependent variable (Burns, 2000). It can be calculated using a number of different methods. In this PhD, two methods were used for calculating effect size. The first, Cohen's  $d$ , was used to calculate the size of the effect in the paired samples t-tests, one way ANOVAs, one way RM-ANOVAs, and MANOVAs. It was calculated by dividing the difference in sample means by the standard deviation of the control group –  $(M_1 - M_2) / SD$ . Cohen's  $f^2$  was calculated to reflect the size of the predictive effect in the multiple regression analysis, using the Adjusted  $R^2$  values:  $f^2 = R^2 / (1 - R^2)$  (Cohen, 1992). Effect size conventions were followed in both cases. For Cohen's  $d$ , a small effect was within the range  $.20 \leq d \leq .49$ , a medium effect was within the range  $.50 \leq d \leq .79$ , and a large effect was  $d \geq .80$  (Cohen, 1988). For the  $f^2$  measure, a trivial effect was  $\leq .02$ , a small effect was  $\leq .15$ , a medium effect was  $< .35$ , and a large effect was  $> .35$  (Cohen, 1992). Both  $d$  and  $f^2$  were chosen because they are well respected measures of effect size (Field, 2013b). Cohen's  $d$  was chosen for most of the analyses because it reflects the size of the effect in standard deviation units, rather than as a percentage (%) of variance, and this was easier to understand and interpret in the context of this PhD thesis.

### **3.5 Summary**

The design of all three evaluation studies in this thesis were informed by the methodological literature reviewed in this chapter. The first two studies are evaluations of the implementation of a programme curriculum, and the final study is an evaluation of the outcomes of Kiwi Can, a social health promotion programme. They each utilise quantitative, survey-style data collection methods, and descriptions

of the design, methods and results for each of the studies can be found in the following three chapters.

## **Chapter 4: The Implementation of the Health Education Curriculum in New Zealand Schools**

### **4.1 Introduction**

To date, little research has been conducted in New Zealand on the delivery of the health education curriculum in primary or intermediate schools. This study was designed to contribute to the knowledge base in this area by exploring the delivery of the health education curriculum from the perspective of health education practitioners. It provides cross-sectional data on the methods and frequency of curriculum delivery, and explores the impact of teacher beliefs and proposed barriers.

The study was conducted following formal approval by the Human Participants Ethics Committee at The University of Auckland (reference number 8362), and was guided by six main research questions –

1. How frequently is health education, as a general area of study, taught in primary and intermediate schools around New Zealand?
2. How frequently do primary and intermediate school teachers deliver lessons across the four health dimensions, four KLAs, and 11 AOs contained in the health education curriculum?
  - 2(i) How frequently are social health-related lessons delivered in primary and intermediate schools?
3. What impact do teacher beliefs about the importance of teaching a health dimension have on the delivery of lessons across the four health dimensions, four KLAs, and 11 AOs?
  - 3(i) What impact do teacher beliefs about the importance of teaching a health dimension have on the delivery of social health-related lessons?
4. What impact do teacher beliefs about the benefit of teaching a health dimension have on the delivery of lessons across the four health dimensions, four KLAs, and 11 AOs?

- 4(i) What impact do teacher beliefs about the benefit of teaching a health dimension have on the delivery of social health-related lessons?
5. What are the most frequently cited barriers to the delivery of the health education curriculum in primary and intermediate schools around New Zealand?
6. Are there statistically significant differences in self-reported delivery of health education as a function of school position, years teaching health education, and/or year level taught?

## **4.2 Study Design**

A single occasion, self-report survey design was used for this quantitative study. The participants were health education practitioners in primary and intermediate schools around New Zealand, who were purposively selected according to a stratified sampling frame (see Chapter 3). The survey data were collected using an anonymous, electronic questionnaire.

### **4.2.1 Participants.**

All primary and intermediate school health education teachers in public, mainstream schools in New Zealand were eligible to participate in this study. They were contacted via their school principal or the president of their health and physical education association. The recruitment methods varied by geographic area. In Auckland, potential participants were invited to participate directly by their school principal. Hardcopy posters were mailed to each of the primary and intermediate schools in the city ( $n = 356$ ), and personalised emails were also sent to principals informing them about the research, with a link to the electronic survey. Potential participants from the rest of the country were invited indirectly via newsletter and/or email correspondence from the presidents of The Physical Education in New Zealand Association, The New Zealand Health Education Association, and the Health Promoting Schools organisation. The presidents forwarded an electronic invitation, along with the survey link, to all their members nationwide (approximately 2,000 health education personnel). The study was also advertised in each organisation's

electronic newsletters. Please see Appendix C for a copy of the Participant Information Sheet for this study.

At the end of the recruitment phase, a total of 214 primary school educators from around New Zealand had agreed to participate in the research project, and 198 completed the survey (a response rate of approximately 10%). Of the 198 participants, just over half (51.4%) were classroom teachers, while the remainder of the sample (48.6%) were either primary school principals, teaching principals, deputy principals, teaching deputy principals, curriculum/syndicate leaders, and/or heads of department (see Table 1). At the time of survey administration, 55.5% of the participants had more than 10 years of experience teaching the health education curriculum, and 40% had more than 15 years of experience. Analysis of the year levels taught by participants revealed a relatively even spread across the spectrum (New Entrant to Year 8), though 17 participants (presumably principals or deputy principals) did not affiliate themselves with a year group. Finally, given the targeted recruitment drive in Auckland, the vast majority of participants (nearly 75%) were from this region.

Table 1

*Frequency Count of Demographic Characteristics for Survey Participants*

Demographic Characteristic	Frequency ( <i>n</i> )	Percent (%)
School Position		
School Leader or Senior Management	51	25.8
Curriculum/Syndicate Leader	33	16.7
Classroom Teacher/Subject Teacher	97	51.4
Other	12	6.1
HE Teaching Experience		
< 2 years	21	10.6
2-5 years	38	19.2
6-10 years	31	15.7
11-15 years	34	17.2
> 15 years	74	37.4
Teaching Level		
New Entrant – Year 3	73	36.9
Year 4 – Year 6	67	33.8
Year 7 – Year 8	41	20.7
Not applicable	17	8.6
School Region		
Auckland	146	73.7
North Island		
Northland	1	.5
Waikato	7	3.5
Bay of Plenty	4	2.0
Gisbourne	2	1.0
Hawke's Bay	1	.5
Taranaki	3	1.5
Manawatu-Wanganui	9	4.5
Wellington	9	4.5
South Island		
Tasman	1	.5
Nelson	1	.5
Canterbury	5	2.5
Otago	3	1.5
Southland	6	3.0
Total	198	100.0

*Note.* HE = health education.

#### 4.2.2 Instrument.

The questionnaire design was informed by a review of health education programme evaluation literature and, most notably, by the content of the Curriculum Stocktake: National School Sampling Survey Health and Physical Education Questionnaire (McGee et al., 2004). This survey was designed in 2000 to collect self-report data on teacher practice in HPE in New Zealand primary and secondary schools (McGee et al., 2004). The original survey was completed by 775 HPE teachers at 291 randomly selected schools across the country. It contained eight demographic questions,

followed by a range of questions about personal teaching practices, including questions about the impact of curriculum structure on practice ( $n = 4$  questions), personal delivery of the curriculum content ( $n = 11$ ), methods of implementation ( $n = 3$ ), use of resources ( $n = 6$ ), and exposure to professional development opportunities ( $n = 4$ ). The questionnaire concluded with questions related to teacher confidence ( $n = 2$ ), inclusivity ( $n = 1$ ), assessment practices ( $n = 4$ ), and barriers to delivery ( $n = 1$ ). At the end of the survey participants were invited to comment on any related issues ( $n = 1$ ). For each of the 45 questions, individuals were required to provide frequency data as well as written comment. See Appendix D for an excerpt from the original stocktake questionnaire.

Of these 45 items, seven were replicated as part of this thesis. They were selected because they pertained to the topics relevant for this study, specifically those related to teacher beliefs about health education, personal implementation practices, and barriers to delivery. An additional 13 questions were designed specifically for inclusion in the questionnaire used for this study. Therefore, the final measurement tool consisted of 19 close-ended questions, and had a total of 56 items. Five of the questions were accompanied by space for written comment. The survey was divided into four sections. In the first section, participants were asked to provide demographic information ( $n = 5$  questions). In the second section participants were asked to indicate how important they felt it was to teach children about each of the dimensions of health – social health (SocH), mental health (MH), emotional health (EH), and physical health (PH), and they were also asked to indicate how influential they believed each dimension was on student wellbeing using a 5-point rating scale from 1 = *very important/influential* to 5 = *not important/influential* ( $n = 4$  questions). In the third section, participants reported on the format and frequency of delivery across the health dimensions, KLAs, and AOs ( $n = 9$  questions) defined in the current health education curriculum (MOE, 2007). Item-response formats in this section varied from fill-in-the-blank questions (e.g., “On average, how many health education classes do you teach over the course of a week?”) to a 5-point rating scale for frequency (1 = *always/very frequently* to 5 = *very rarely/never*). In section four, participants were asked to rate the degree to which they felt that six specific barriers impeded their delivery of health education content ( $n = 1$  question). For this section, participants used a 5-point rating scale (1 = *very strongly affected my practice*, to 5 = *did not affect*

*my practice at all*). Questions about teacher beliefs and barriers were included in the survey to shed light on the nature and impact of factors that influenced teacher practice in health education. Please see Appendix E for a copy of the final questionnaire.

### **4.3 Data Collection**

The questionnaire was deployed on the online, web-based survey programme, Survey Monkey. Four individuals trialled the questionnaire (excluded from results), and format modifications were made that integrated their feedback. The data collection period lasted for six months (February 2013 to July 2013). Questionnaire completion took approximately 15 minutes. At the end of the six-month data collection period, the data ( $n = 214$  surveys) was downloaded and deleted from the Survey Monkey platform.

The questionnaire was screened for missing data, and as the survey had 56 items, 16 surveys with more than five incomplete or missing items ( $> 10\%$  missing data) were immediately removed from the study. An additional 11 surveys, with between one and four missing responses, were imputed with the EM technique, since the MCAR test indicated that the missing responses met that criterion,  $\chi^2 = 1363.47$ ,  $df = 1228$ ,  $p = .40$ . This produced a final data set of 198 complete questionnaires.

### **4.4 Data Analysis**

The questions were analysed using a variety of conventional techniques. Descriptive statistics for each item ( $M$ ,  $SD$ , kurtosis, skewness, etc.) were generated and used to assess the normality of each variable. The internal reliability of the three pre-defined scales (Teacher Beliefs, Delivery of Curriculum, and Barriers to Delivery), was also tested using Cronbach alphas. Results indicated that the three scales, Teacher Beliefs ( $\alpha = .89$ ), Delivery of Curriculum ( $\alpha = .82$ ) and Barriers to Delivery ( $\alpha = .81$ ), had moderate to high internal reliability.

At that point, MANOVAs (with main effects and two-way interactions), bivariate correlations, paired samples t-tests, and simultaneous (blockwise) multiple regression analyses, were used to identify statistically significant relationships between personal

beliefs, perceived barriers, and curriculum delivery. They were also used to explore differences in participant responses across three main demographic categories – position in the school, years of teaching health education, and year level taught. Given the large discrepancy in participant numbers between regions, it was decided not to compare the data across regions.

The open-ended comment data were thematically coded using a deductive content-based approach. The thematic codes pertained to barriers ( $n = 1$  question), frequency of delivery ( $n = 3$  questions), and the format of delivery ( $n = 1$  question). All comments with similar content were grouped into a single theme and assigned a single code (e.g., Lack of Resources, Frequency of Delivery, Sexuality Education, etc). At the end of the coding process, two of the five categories (Frequency of KLAs – School, and Frequency of KLAs – Personal) were merged because the content was conceptually very similar and featured a high degree of overlap and repetition.

## **4.5 Results**

The results of this study are presented in three parts. In Part one, the results of the open-ended coding process are described briefly. In Part two, data on the format and frequency of delivery in health education is presented. In Part three, descriptive and analytic data are presented on the prevalence and frequency with which barriers impeded successful delivery of the health education curriculum.

### **4.5.1 The open-ended comment data.**

At the end of the deductive coding process, four discrete categories of data were created. They were as follows: (i) Barriers to Delivery ( $n = 42$  comments); (ii) Frequency of Delivery – KLAs ( $n = 35$  comments); (iii) Frequency of Delivery – AOs ( $n = 12$  comments); and (iv) the Use of External Providers ( $n = 61$  comments). The data generated through this open-ended coding process is presented alongside the self-report rating data in the appropriate thematically related part of the results. The open-ended coding is used to illustrate and/or further elaborate on the findings from the self-report rating analysis.

## 4.5.2 The delivery of the health education curriculum.

### 4.5.2.1 *Format of health education delivery.*

According to approximately half of the survey participants ( $n = 99$ ), health education was delivered as an integrated area of study. Integrated lessons involved teaching health education content in the context of another learning area, and not as a stand-alone subject. When it was delivered as a stand-alone subject it was most frequently taught by teachers who had been in the field for six to 15 years (the *mid-range* teaching group). These individuals reported that, on average, each stand-alone health education class was 20 to 30 minutes in length ( $M = 26.81$  minutes,  $SD = 19.24$ ). They also reported delivering one to two classes per week ( $M = 1.61$ ,  $SD = 1.92$ ), and although there was variation by year level, the differences were not statistically significant,  $\chi^2 = 85.13$ ,  $df = 72$ ,  $p = .14$ .

About one third ( $n = 61$ ) of participants also reported the use of external providers for health education delivery. External providers were individuals and/or community-based organisations commissioned to provide an educational service, and were not permanently contracted members of the teaching staff. All the external providers used delivered the entire health education lesson, and the teacher was a passive observer. Participants also reported that most were from large, national organisations, and that their employees did not have formal teaching qualifications or a working knowledge of the New Zealand health education curriculum. The most commonly cited external providers were the Life Education Caravan ( $n = 15$ ), the New Zealand police ( $n = 10$ ), and public health nurses ( $n = 6$ ). Programming primarily focused on physical health promotion, disease prevention, and safety management. Only two of the external providers, Kiwi Can ( $n = 1$ ), and Cool Bananas ( $n = 1$ ), delivered educational programming in the emotional health, mental health, and social health dimensions (in both cases, all three). Neither of these providers delivered programming in the physical health area.

#### 4.5.2.2 Frequency of delivery of health dimensions.

The health education lessons delivered by the teachers in this study focused most frequently on physical health, followed by emotional health, mental health, and social health (see Table 2). As presented in Table 2, higher mean scores indicate lower frequency.

Table 2

*Self-Reported Frequency of Delivery Across the Four Health Dimensions*

Dimension	<i>M</i>	<i>SD</i>	Mean Score Difference Measures [Paired <i>t</i> -test (Cohen's <i>d</i> )]			
			1.	2.	3.	4.
1. PH	2.30	1.02	—			
2. EH	2.59	1.10	-3.84*** (.26)	—		
3. MH	2.62	1.10	-9.04*** (.67)	-8.65***(.62)	—	
4. SH	3.04	1.17	-4.39*** (.29)	.64	-6.61*** (.45)	—

*Note.* Mean values based on a 5-point rating scale (1 = *very frequently*, 2 = *frequently*, 3 = *occasionally*, 4 = *seldom*, 5 = *very rarely/never*); PH = physical health; EH = emotional health; MH = mental health; SH = social health.

\*\*\* =  $p < .001$ .

There were both statistically and practically significant differences among the self-reported delivery of health areas. Physical health programming was offered substantially more often ( $d > .60$ ) than mental health, and a small amount more than emotional and social health. Emotional health was delivered substantially more often than mental health ( $d > .60$ ), while mental health was delivered more often than social health. There were no statistically significant differences in the frequency of delivery between emotional and social health ( $p \geq .05$ ).

When the frequency of delivery was compared between participant groups, the results of the MANOVAs (using main effects and all two-way interactions) showed no statistically significant differences for position in school,  $V = .05$ ;  $F_{(12, 441)} = .60$ ,  $p = .84$ ), years teaching health education,  $V = .11$ ;  $F_{(16, 592)} = 1.05$ ,  $p = .40$ , or year level taught,  $V = .12$ ;  $F_{(12, 441)} = 1.49$ ,  $p = .13$ . Hence, the practitioners, regardless of position, experience teaching health education, or year level taught, were consistent in their practices. They delivered physical health-related lessons most frequently, and social health-related lessons least frequently.

#### 4.5.2.3 Frequency of delivery of KLAs and AOs.

In this section, the frequency of teaching health education classes related to the four KLAs and the 11 AOs contained in the New Zealand HPE curriculum is examined. Table 3 shows that Body Care, and Physical Safety and Food and Nutrition, the two physical health-related KLAs, had very similar and generally high frequencies, with small effect size differences. However, there were large differences ( $d > .60$ ) between these two topics and Mental Health and Sexuality Education, which were delivered much less frequently. The difference between Mental Health and Sexuality Education was small, with the former delivered somewhat more frequently.

Table 3

*Means, Standard Deviations, and Output from Paired Samples t-tests for the KLAs in the New Zealand Health Education Curriculum*

Key Learning Area	M	SD	Mean Score Difference Measures [Paired t-test Cohen's d]			
			1.	2.	3.	4.
1. Body Care and Physical Safety	2.10	.76	—			
2. Food and Nutrition	2.22	.79	-2.80** (.15)	—		
3. Mental Health	3.10	.95	-15.04*** (1.05)	-13.62*** (.93)	—	
4. Sexuality Education	3.43	.98	-16.66*** (1.36)	-16.43*** (1.23)	-4.24*** (.34)	—

*Note.* Mean values based on a 5-point rating scale (1 = *very frequently*, 2 = *frequently*, 3 = *occasionally*, 4 = *seldom*, 5 = *very rarely/never*).

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ .

In some cases, the KLA – Sexuality Education was not delivered at all. This was evidenced by open-ended comments:

I'm not really sure about the last one (Sexuality Education) – I think it's more at the senior end of the school. (Par134)

I don't think mental or sexuality education are taught very often at all. Not as units (at a junior level anyway). Individual teachers may touch on those topics with their class briefly, during the day-to-day lessons though. (Par45)

Sexuality Education [is] on the agenda this year, as I have not covered it. I have been following our three-year plan for Health (and we haven't had any sex education). (Par116)

I teach food and biotechnology, so sexuality education is not part of the programme. (Par99)

We only teach about puberty, not about sex education. (Par34)

Sexuality isn't something we teach. Again other topics are frequently addressed outside of a teaching lesson. (Par76)

We formally teach Keeping Ourselves Safe (KOS) every 4 years or so – however I teach a single sex class, so teaching sex education for safety or cleanliness arises occasionally. (Par155)

While KLA delivery was infrequent, and in some cases non-existent, AO delivery was not. In fact, programming for most AOs was offered frequently to very frequently. The AOs delivered with the greatest frequency were in the mental and emotional health dimensions (i.e., Identity, Sensitivity and Respect, and Positive Attitudes). These were followed by the delivery of AOs in the individual social health dimension. Individual social health AOs are associated with learning about social adjustment, social skills development, and social cohesion (i.e., Relationships, Interpersonal Skills, and Social Attitudes and Values). They differ from socioecological health AOs that focus on learning about the wider community and society (i.e., Community Resources, Rights, Responsibilities and Laws, and People and the Environment). In this study, the socioecological AOs were delivered significantly less frequently than all other AOs. The size of the effect was large for the Community Resources AO ( $d > .80$ ), and ranged from small to large for the Rights, Responsibilities and Laws AO ( $.41 \leq d \leq 1.44$ ), and from trivial to large for the People and Environment AO ( $.18 \leq d \leq 1.11$ ). See Table 4 for the results of paired t-tests for frequency of AO delivery.

When multiple analysis of variance tests (using main effects and all two-way interactions) were used to explore differences in the frequency of AO and KLA delivery, there were no statistically significant differences for position,  $V = .14$ ;

$F_{(12,441)} = 1.78, p = .06$ , years teaching health education,  $V = .10$ ;  $F_{(16, 592)} = .96, p = .50$ , or year level taught,  $V = .05$ ;  $F_{(12, 441)} = .61, p = .84$  in KLA frequency, and no statistically significant differences in the frequency of delivery for any of the AOs by position,  $V = .20$ ;  $F_{(33, 420)} = .90, p = .63$ , years teaching health education,  $V = .26$ ;  $F_{(44, 564)} = .90, p = .66$ , or year level taught,  $V = .15$ ;  $F_{(33, 420)} = .66, p = .93$ . This indicates that all KLAs and AOs were delivered with similar frequency across the participants.

Table 4

Means, Standard Deviations, and Output from Paired Samples *t*-tests for Delivery of the Achievement Objectives (AOs) in the New Zealand HPE Curriculum

Achievement Objective	Mean Score Difference Measures [Paired <i>t</i> -test (Cohen's <i>d</i> )]												
	<i>M</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
<b>1. Personal Growth and Development</b>	2.20	.76	—										
<b>2. Safety Management</b>	2.18	.79	.55	—									
<b>3. Personal Identity</b>	2.08	.75	2.39* (.16)	1.86	—								
<b>4. Identity, Sensitivity and Respect</b>	1.66	.61	5.60*** (.71)	5.31*** (.66)	4.80*** (.56)	—							
<b>5. Positive Attitudes</b>	1.74	.68	10.04*** (.61)	9.98*** (.56)	9.20*** (.49)	5.33*** (.13)	—						
<b>6. Relationships</b>	1.87	.71	8.26*** (.43)	8.09*** (.39)	6.83*** (.28)	3.33** (.34)	-2.97** (.19)	—					
<b>7. Interpersonal Skills</b>	1.81	.73	6.68*** (.51)	6.51*** (.47)	5.54*** (.36)	1.76 (.25)	-3.52** (.10)	-1.84	—				
<b>8. Social Attitudes and Values</b>	1.89	.75	4.80*** (.41)	4.47*** (.37)	3.40** (.25)	-.65	-4.79*** (.22)	-3.33** (.03)	-2.03* (.11)	—			
<b>9. Community Resources</b>	2.88	.83	-11.28*** (.89)	-11.09*** (.89)	-13.33*** (1.07)	-17.19*** (2.00)	-20.25*** (1.68)	-18.48*** (1.42)	-16.92*** (1.47)	-17.15*** (1.32)	—		
<b>10. Rights, Responsibilities and Laws</b>	2.54	.87	-5.71*** (.45)	-5.86*** (.46)	-6.99*** (.61)	-10.70*** (1.44)	-14.20*** (1.18)	-13.48*** (.94)	-11.51*** (1.00)	-10.73*** (.87)	5.62*** (.41)	—	
<b>11. People and the Environment</b>	2.34	.84	-2.29* (.18)	2.73** (.20)	-4.29*** (.35)	-8.21*** (1.11)	-11.66*** (.88)	-10.08*** (.66)	-8.55*** (.73)	-7.71*** (.60)	9.52*** (.65)	3.84*** (.23)	—

Note. Mean values based on a 5-point rating scale (1 = *very frequently*, 2 = *frequently*, 3 = *occasionally*, 4 = *seldom*, 5 = *very rarely/never*); values in blue are related to the physical health area, values in purple are related to mental and emotional health, values in red are related to individual social health and values in orange are associated with learning in the socioecological social health dimension.

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ .

It must be noted that participants in this study reported integrating teaching and learning in the mental, emotional, and social health-related KLAs and AOs with other subject areas. Some also reported teaching mental, emotional, and/or social health-related health education during a single week (block) of the school year only. This is evidenced by the following quotes:

These (AOs) are not taught as 'subjects', but entwined with my teaching. For example – personal growth students are responsible for their own actions, show respect, work on their educational progressions. (Par144)

These topics often are integrated into an enquiry topic for the term. (Par55)

We integrate social health into our other subject areas. (Par2)

In a few cases health education was actually taught. In this case, it was integrated into a week-long block (i.e., a school camp). (Par121)

Overall, it appears that health education practitioners are delivering, on average, one short health education lesson per week. The content of those lessons is most frequently in the physical health dimension, and least frequently in the social health dimension. This pattern holds for the delivery of lessons related to the KLA, but not for the AOs. Within the AO space, mental, emotional, and individual social health lessons are delivered most frequently. However, they are rarely taught as stand-alone lessons. Instead, they are frequently integrated with learning in other areas and/or taught as a single block of lessons.

In order to identify possible explanations for the discrepancy in frequency and method of delivery between dimensions, teacher beliefs concerning barriers were explored as possible confounding, mitigating, and/or causal factors.

### 4.5.3 Teacher beliefs.

For this section, personal beliefs were sub-divided into two areas – beliefs about the importance of teaching a particular dimension of health (*Importance*), and beliefs about the positive effect that teaching and learning in the area might have on student wellbeing (*Benefit*).

#### 4.5.3.1 Importance.

Survey participants reported that it was important and/or very important to deliver health education programming to students in each of the four dimensions of health – physical health, emotional health, mental health, and social health (see Table 5). Of the four, teaching physical health was rated as being the most important, followed by emotional health, mental health, and social health. According to the paired t-test output, participants believed that teaching physical health and emotional health were significantly more important than teaching children about mental and/or social health. The size of the effect was small in each case ( $.15 \geq d \leq .47$ ).

Table 5

*Means, Standard Deviations, Confidence Intervals, and Output of Paired Samples t-tests for Importance of Teaching the Four Dimensions of Health on Student Wellbeing*

Importance	<i>M</i>	<i>SD</i>	95% CI		Mean Score Difference Measures [Paired <i>t</i> -test (Cohen's <i>d</i> )]			
			<i>LL</i>	<i>UL</i>	1.	2.	3.	4.
1. PH	1.20	.43	1.14	1.26	—			
2. EH	1.27	.55	1.19	1.35	-1.77	—		
3. MH	1.35	.56	1.28	1.43	-4.22*** (.35)	-3.16** (.15)	—	
4. SH	1.40	.63	1.32	1.50	-4.64*** (.47)	-3.93*** (.24)	-1.25	—

*Note.* Mean values based on a 5-point rating scale (1 = *very important*, 2 = *important*, 3 = *moderately important*, 4 = *slightly important*, 5 = *not important*); CI= confidence interval; LL= lower limit, UL= upper limit; PH = physical health, EH = emotional health, MH = mental health, SH = social health. \*\*\* =  $p < .001$ ; \*\* =  $p < .01$ .

When mean differences in beliefs were explored using a multiple analysis of variance (using main effects and all two-way interactions), significant differences emerged as a function of years of teaching experience,  $V = .22$ ,  $F_{(16, 592)} = 2.10$ ,  $p = .01$ . Output

from the Games-Howell post hoc, indicated that teachers with six to 10 years of experience reported a stronger belief in the importance of teaching social health than teachers with 2 to 5 years of experience  $V = -.63$ , 95% CI [.08, 1.17],  $p = .01$ . There were no significant differences in beliefs about importance of teaching by position in school,  $V = .13$ ;  $F_{(12, 441)} = 1.65$ ,  $p = .08$ , or by year level taught,  $V = .08$ ;  $F_{(12, 441)} = 1.04$ ,  $p = .41$ . These results suggest that personal beliefs had very little impact on teaching practice in this sample.

Delivery in each of the four health education dimensions was then investigated as a function of personal beliefs. The relationship between delivery and beliefs was explored using bivariate correlations and simultaneous multiple regression analyses. Participant's beliefs about the importance of teaching the four health dimensions were moderately correlated ( $r = .19$  to  $.39$ ) with delivery (see Table 6). Delivery in each dimension was most strongly correlated with personal beliefs about the importance of delivering lessons in the corresponding health dimension.

Table 6

*Inter-correlations for Importance and Frequency of Delivery Across the Health Dimensions*

Frequency	Importance			
	1.	2.	3.	4.
1. Physical health	<b>0.21**</b>	0.39**	0.23**	0.28**
2. Emotional health	0.21**	<b>0.31**</b>	0.29**	0.29**
3. Mental health	0.19**	0.26**	<b>0.39**</b>	0.25**
4. Social health	0.21**	0.26**	0.26**	<b>0.35**</b>

*Note.* Values in bold are matched health areas.

\*\* =  $p < .01$ .

Multiple regression models showed that importance beliefs significantly explained the self-reported frequency of teaching in the mental and social, but not the emotional or physical health dimensions (see Table 7). The frequency of physical health delivery was positively explained, albeit weakly ( $f^2 < .15$ ), by beliefs about the importance of emotional health, and negatively by beliefs about the importance of mental health and social health. The frequency of mental and social health delivery was moderately explained by beliefs about importance ( $f^2 = .14$  to  $.16$ ). Importance beliefs did not explain delivery across any of the KLAs. The output from the bivariate correlational matrix (see Table 8) indicated that beliefs about the importance of teaching health

dimensions and the delivery of KLAs ranged from indistinguishable to weak ( $r = .21$ ). These zero to weak correlations translated into statistically non-significant regressions from importance of health dimensions to frequency of KLAs. In contrast, the importance beliefs had more influence on the reported delivery of AOs. For Personal Identity and Personal Growth and Development, there were statistically significant correlations between beliefs about the importance of all four health dimensions and practice. The correlations were moderate in size, ranging from  $r = .15$  to  $r = .29$ . Significant moderate correlations were also evident between beliefs about the importance of emotional and mental health, and delivery in this area (Positive Attitudes and Identity, Sensitivity and Respect). There were no statistically significant correlations between beliefs about the importance of social health and the frequency of delivery of AOs in this dimension. Instead, delivery in the individual social health AOs was more strongly associated with beliefs about the importance of teaching emotional and mental health (see Table 9).

Table 7

*Multiple Regression Output for Personal Beliefs About the Importance of a Health Dimension as a Predictor of Delivery in Health Areas*

Predictor	Dependent Variable	Standardized Regression Weight ( $\beta$ )	95% CI		$p$	Adjusted $R^2$	Effect size ( $f^2$ )
			LL	UL			
ImpSocH	FreqSocH	.35	.31	.95	***	.14	.16
ImpMH	FreqMH	.27	.15	.95	**	.12	.14
ImpEH	FreqPH	-.28	-.87	-.11	**	.08	.09
ImpMH	FreqPH	.26	.07	.83	*	.08	.09
ImpSocH	FreqPH	.24	.11	.70	**	.08	.09

*Note.* ImpPH=importance of teaching physical health; ImpEH= importance of teaching emotional health, ImpMH=importance of teaching mental health; ImpSocH=importance of teaching social health; Freq=self-reported score for frequency of delivery across each health dimension. CI = confidence interval; LL=lower limit, UL = upper limit.

\*\*\* =  $p < .001$ ; \*\* =  $p < .01$ ; \* =  $p < .05$ .

Table 8

*Inter-correlations for Importance and Frequency of Delivery of KLAs*

Frequency – KLAs	Importance			
	PH	EH	MH	SH
Body Care and Personal Safety	<b>.04</b>	-.03	.06	-.08
Food and Nutrition	<b>.12</b>	.10	.14	.10
Mental Health	-.05	.03	<b>.11</b>	.00
Sexuality Education	<b>.02</b>	<b>.21**</b>	<b>.19**</b>	.17*

*Note.* Values in bold are matched health constructs; PH = physical health, EH = emotional health, MH = mental health, SH = social health.

\*\* =  $p < .01$ , \* =  $p < .05$ .

Table 9

*Inter-correlations for Importance and Frequency of Delivery of AOs*

Frequency – AOs	Importance			
	Physical health	Emotional health	Mental health	Social health
Personal Growth and Development	<b>.11*</b>	.23**	.28**	.15*
Safety Management	<b>.09</b>	.09	.17*	.12
Personal Identity	.16*	<b>.22**</b>	<b>.29**</b>	.21**
Positive Attitudes	.10	<b>.25**</b>	<b>.26**</b>	.13
Relationships	.06	.23**	.31**	<b>.14</b>
Identity, Sensitivity and Respect	.12	<b>.23*</b>	<b>.21**</b>	<b>.12</b>
Interpersonal Skills	.08	.22**	.29**	<b>.07</b>
Social Attitudes and Values	.10	.04	.13	<b>.07</b>
Community Resources	.05	.00	.13	<b>.02</b>
Rights, Responsibilities and Laws	.02	.08	.12	<b>.05</b>
People and the Environment	.11	.07	.11	<b>.09</b>

*Note.* Values in bold are matched health constructs.

\*\* =  $p < .01$ , \* =  $p < .05$ .

When self-reported scores for importance of teaching health concepts were used to explain the frequency of delivery across the 11 AOs, three statistically significant relationships emerged. Beliefs about the importance of teaching mental health were associated with a reduction in the frequency of delivery for Interpersonal Skills,  $\beta = .38$ , 95% CI [.20, .79],  $p < .01$ , and Relationships,  $\beta = .42$ , 95% CI [.23, .78],  $p = .03$ , an beliefs about the importance of teaching Social Health positively explained delivery in Social Attitudes and Values,  $\beta = -.22$ , 95% CI [-.51, -.03],  $p = .03$ . The size of the effect was trivial in each case (Adjusted  $R^2 = < .10$ ;  $f^2 = < .10$ ). There were no statistically significant correlations or regression effects for beliefs and socioecological health dimensions.

Overall, the results suggest that health education teacher practice (delivery) was not strongly influenced by personal beliefs about the importance of teaching in the dimension. This was most notable in the physical, emotional, and social health dimensions. Within the social health dimension, and social health AOs specifically, delivery was most closely associated with beliefs about the importance of mental and emotional health. This suggests that health education teachers may not fully understand the social health area, may not have strong beliefs about its importance, or may not have chosen to use their beliefs about the importance of the area to guide their teaching.

#### 4.5.3.2 Benefit.

After being queried about importance beliefs, participants were then asked about the perceived benefit of teaching and learning in health education for children’s health and wellbeing. According to the teachers in this sample, teaching children about emotional health was most beneficial, followed by teaching them about physical health, mental health, and social health. Output from paired samples t-tests indicated that participants believed that teaching children about emotional health was significantly more beneficial than teaching them about physical, mental, and/or social health. The effect sizes ranged from very small ( $d = .16$ ) to moderate ( $d = .45$ ). Participants also felt that teaching children about physical health was slightly more beneficial for their wellbeing than teaching them about social health. There were no statistically significant differences in self-reported benefit for mental health (see Table 10).

Table 10

*Means, Standard Deviations, Confidence Intervals, and Output of Paired Samples t-tests for Teaching the Four Dimensions of Health on Student Wellbeing*

Benefit	<i>M</i>	<i>SD</i>	95% CI		Mean Score Difference Measures [Paired <i>t</i> -test (Cohen’s <i>d</i> )]			
			<i>LL</i>	<i>UL</i>	1.	2.	3.	4.
1. PH	1.37	.56	1.29	1.45	—			
2. EH	1.28	.49	1.21	1.35	2.39* (.16)	—		
3. MH	1.44	.69	1.36	1.55	-1.60	-4.67*** (.33)	—	
4. SH	1.50	.63	1.41	1.59	-3.10** (.23)	-5.69*** (.45)	-1.17	—

*Note.* Mean values based on a 5-point rating scale (1 = *very influential*, 2 = *influential*, 3 = *moderately influential*, 4 = *slightly influential*, 5 = *not influential*); CI = confidence interval; LL=lower limit, UL = upper limit, PH = physical health, EH = emotional health, MH = mental health, SH = social health. \*\*\* =  $p < .001$ ; \* =  $p < .05$ .

When mean differences in beliefs about benefit were explored using a multiple analysis of variance (using main effects and all two-way interactions), Pillai’s trace output showed a statistically significant difference as a function of position,  $V = .14$ ,  $F_{(12, 441)} = 1.85$ ,  $p = .04$ . Output from the Games-Howell post hoc, however, did not

support this finding ( $p \geq .05$ ). There were no statistically significant differences in beliefs about the benefit of teaching the health dimensions as a function of either years teaching health education,  $V = .13$ ,  $F_{(16, 592)} = 1.23$ ,  $p = .24$ , or year level taught,  $V = .08$ ;  $F_{(12, 441)} = .96$ ,  $p = .49$ . This suggests that personal beliefs about the benefit of teaching the health dimensions are similar across participants.

Participants' beliefs about the benefit of delivery were positively but weakly to moderately associated with practice across three of the four dimensions. There was no significant correlation between the benefit of teaching physical health, and delivery in this area (see Table 11). When the matched diagonal correlations were compared, the within trait correlations were higher than the between trait (off-diagonal) correlations for three of the four dimensions – physical, social and mental health. Beliefs about the benefit of teaching emotional health were not as strongly correlated with delivery in the same dimension. In fact, beliefs about the benefit of teaching emotional health were more strongly correlated with practice in the mental health dimension. This suggests that there is some conceptual overlap between the emotional health and mental health constructs. The teachers may feel that children can benefit both emotionally and mentally from participation in lessons that are designated as either mental and/or emotional health-related.

Table 11

*Inter-correlations for Benefit and Frequency of Delivery Across the Health Dimensions*

Frequency	Benefit of Health Dimensions			
	1.	2.	3.	4.
1. Physical health	<b>.11</b>	.12	.23**	.17*
2. Emotional health	.06	<b>.20**</b>	.20**	.17*
3. Social health	.09	.15*	<b>.36**</b>	.18*
4. Mental health	.09	.24**	.24**	<b>.28**</b>

*Note.* Corresponding constructs have been highlighted in bold.

\*\* =  $p < .01$ , \* =  $p < .05$ .

When beliefs about influence were explored as predictors of delivery, two significant relationships emerged. In this sample, beliefs about the benefit of teaching mental health,  $\beta = -.20$ , 95% CI [.03, .43],  $p = .02$ , and social health  $\beta = -.39$ , 95% CI [.35, -1.02],  $p < .01$  significantly explained the frequency of delivery in the same dimension. The magnitude of influence was small for both social health (Adjusted  $R^2 = .13$ ,  $f^2 =$

.15) and mental health (Adjusted  $R^2 = .10$ ;  $f^2 = .11$ ). There were no statistically significant effects for beliefs about the benefit of physical health or emotional health on frequency of delivery.

The relationships between beliefs about benefit and delivery in the KLAs and AOs produced similarly weak results. In fact, the Pearson correlation co-efficients indicated that beliefs about benefit and KLA delivery were negligible at best ( $r = .00 - .16$ ), while the simultaneous multiple regression analysis did not produce a single significant regression pathway between benefit and KLA delivery. The patterns changed somewhat for the AO analysis. In this case, beliefs about the benefit of teaching the health dimensions were statistically significantly correlated with delivery across nine of the eleven achievement objectives (see Table 12). They were not however, correlated with the frequency of delivery in the same dimension. In fact, beliefs about the importance of teaching physical health were most strongly correlated with delivery in the social health-related AOs, as were beliefs about teaching emotional and mental health. Beliefs about the benefits of teaching social health were most strongly correlated with delivery in the physical health-oriented AOs (Personal Growth, and Development and Safety Management). Output from a simultaneous multiple regression analysis indicated that there were no statistically significant effects for any of these relationships (Adjusted  $R^2 \leq .05$ ;  $f^2 \leq .05$ ).

Table 12

*Inter-correlations for Benefit and Frequency of Delivery in AOs*

Frequency	Benefit of Health Dimensions			
	Physical health	Emotional health	Mental health	Social health
Personal Growth and Development	<b>.07</b>	.13	.11	.16*
Safety Management	<b>.13</b>	.10	.07	.20*
Personal Identity	.08	<b>.05</b>	<b>.02</b>	<b>.13</b>
Positive Attitudes	.18*	<b>.10</b>	<b>.13</b>	.16*
Relationships	.20*	.16*	.22*	<b>.21**</b>
Identity, Sensitivity and Respect	.24*	<b>.12</b>	<b>.12</b>	<b>.11</b>
Interpersonal Skills	.16*	.18*	.20*	<b>.15*</b>
Social Attitudes and Values	.15*	.08	.11	<b>.02</b>
Community Resources	.05	.05	.05	<b>.11</b>
Rights, Responsibilities and Laws	.10	-.01	.10	<b>.11</b>
People and the Environment	.10	-.04	-.01	<b>.11</b>

*Note.* Corresponding constructs have been highlighted in bold.

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ , \* =  $p < .05$ .

In general, beliefs about the influence that teaching health education had on well-being were weakly associated ( $r \leq .30$ ) with practice across the mental, emotional, and social health dimensions. However, they were generally not, except in a few cases, statistically regressed onto the frequency with which the teachers delivered content related to the KLAs or AOs. This indicates that teacher practice (delivery) was not strongly driven by personal beliefs. It also suggests that delivery in health education may be influenced by other internal characteristics of the individual, and/or by factors that are external.

#### 4.5.4 Barriers to delivery.

In the final section of the health education delivery survey, three internal and three external barriers to health education delivery were rated by the participants. The two strongest barriers were both external. They were a perceived lack of time and a lack of resources. Effect sizes between these two barriers and all other barriers were large ( $d > .60$ ). See Table 13 for the output from the paired samples t-tests for barriers.

Table 13

*Means, Standard Deviations, and Paired Samples t-test Results for Self-reported Barriers by Primary and Intermediate School Health Education Practitioners*

Barriers	<i>M</i>	<i>SD</i>	Mean Score Difference Measures [Paired <i>t</i> -test (Cohen's <i>d</i> )						
			1.	2.	3.	4.	5.	6.	
1. Lack of time	2.28	1.06	—						
2. Lack of resources	3.45	1.09	-14.68*** (1.07)	—					
3. Lack of personal knowledge	3.78	1.00	-16.78 *** (1.50)	-4.47*** (1.95)	—				
4. Lack of interest from students	4.18	.91	-21.71*** (1.50)	-9.36*** (.44)	-5.78*** (.44)	—			
5. Lack of personal interest	4.26	.92	-22.52*** (2.09)	-10.07*** (.88)	-7.33*** (.52)	-1.73*** (.09)	—		
6. Lack of support from school management	4.23	1.00	-23.39*** (2.15)	-10.36*** (.78)	-5.41*** (.45)	-.70	.42	—	

*Note.* Mean values based on a 5-point rating scale (1 = very strongly affected my practice, 2 = strongly affected my practice, 3 = somewhat strongly affected my practice, 4 = had a weak effect on my practice, 5 = did not affect my practice at all).

\*\*\* =  $p < .001$ .

In the comments section, participants elaborated on the effect that a lack of time had on their practice. Two examples were:

And, although I hate to admit this, if something needs to be squeezed out of my week due to other curriculum pressures, Health is one of the first areas to go. (Par54)

Just that the curriculum is so crowded, and with the narrowing of delivery in response to National Standards, it's quite common for health to get squeezed out. (Par188)

Two different aspects of lack of resourcing were also specifically touched on:

There is a lack of funding. (Par31)

We struggle to identify and use a good selection of mental health resources. (Par16).

According to the results of a MANOVA using main effects and two-way interactions, there were no statistically significant differences reported for barriers as a function of school position,  $V = .13$ ;  $F_{(18, 435)} = 1.05$ ,  $p = .40$ , years teaching health education,  $V = .19$ ;  $F_{(24, 584)} = 1.20$ ,  $p = .23$ , or year level taught,  $V = .16$ ;  $F_{(18, 435)} = 1.38$ ,  $p = .14$ . This suggests that lack of time and lack of resources were the most significant barriers faced by health education practitioners around the country.

#### ***4.5.4.1 Barriers and frequency of delivery in health dimensions.***

A simultaneous multiple regression analysis showed that three of the six barriers could be used to explain the frequency of delivery in three health dimensions – physical, mental, and social health (see Table 14), although effects were small in scale. Lack of time was significantly associated with increased frequency of delivery in the physical health dimension. This suggests that when time is limited, teachers are significantly more likely to teach lessons in physical health, and significantly less likely to teach lessons in mental health or social health. They are more likely to teach lessons in mental health and in social health however, when they feel that children

lack an interest in the mental and social health areas. This suggests that teachers may interpret children’s lack of interest in social and/or mental health as a sign that they need to receive additional instruction in these dimensions. When teachers personally lack interest in the social health dimension however, they are less likely to deliver social health-related lessons.

Table 14

*Multiple Regression Output for Predictive Effect of Barriers on the Frequency of Delivery Across Health Dimensions*

Barrier	Health Dimension	Standardised Regression Weight ( $\beta$ )	Adjusted $R^2$	Effect Size ( $f^2$ )
Lack of time	Physical	.25**	.03	.03
Lack of time	Mental	-.22*	.04	.04
Lack of personal interest in subject	Social	-.31**	.03	.03
Lack of interest from students	Mental	.38*	.04	.04
Lack of interest from students	Social	.34*	.03	.03

\*\* =  $p < .01$ ; \* =  $p < .05$ .

**4.5.4.2 Barriers and frequency of delivery in KLAs and AOs.**

The frequency of delivery in each KLA was analysed as a function of all six barriers. Only delivery in mental health was significantly predicted by the presence of barriers. It increased (statistically significantly, but small in effect) as a function of a lack of interest from children, and decreased as a function of lack of resources (see Table 15). Teachers who perceived that children were not interested in mental health were more likely to deliver lessons in that subject area, but were less likely to do so if they lacked the resources.

A total of seven AOs were influenced by two different barriers. Lack of personal interest in the subject area resulted in decreases in five AOs (Personal Identity, Positive Attitudes, Interpersonal Skills, Social Attitudes and Values, and People and the Environment), although effects were small. Children’s lack of interest resulted in increased delivery of Body Care and Safety, and Food and Nutrition (though again

with small effect sizes). See Table 15 for the output from the simultaneous multiple regression analyses.

Table 15

*Multiple Regression Output for Predictive Effect of Self-reported Barriers and Frequency of Delivery Across KLAs and AOs*

KLA or AO	Standardised Regression Weight ( $\beta$ )	Adjusted $R^2$	Effect size ( $f^2$ )
Lack of interest - Students			
Mental Health	.31**	.04	.04
Body Care and Safety	.24**	.08	.09
Food and Nutrition	.23**	.04	.04
Lack of resources			
Mental Health	-.23**	.02	.02
Lack of personal interest			
Personal identity	-.25**	.03	.03
Positive attitudes	-.27**	.04	.04
Interpersonal skills	-.21*	.08	.09
Social attitudes and values	-.26**	.07	.08
People and Environment	-.25**	.02	.02

\*\* =  $p < .01$ ; \* =  $p < .05$ .

In general, self-reported delivery in health education was most strongly affected by scores for two external barriers – lack of time and lack of resources, and more weakly affected by a lack of personal interest in the subject area and a lack of interest from children. While a lack of time, resourcing, and personal interest all explained a reduction in frequency of delivery, a lack of interest from children did not.

#### 4.6 Conclusion

The results of this survey study on health education provision in New Zealand primary and intermediate schools showed that a comprehensive health education curriculum was not being offered in, at least, half of the schools. Health education teachers reported delivering a narrow programme focused predominantly on the physical health dimension, with little attention paid to the emotional, mental, and/or social health dimensions. In fact, mental and social health-related classes were delivered significantly less frequently than both emotional and physical health classes ( $p < .05$ ). In this study, social health lessons were reportedly delivered with the least frequency.

The delivery of health education lessons, particularly in the social dimension, were not significantly explained by personal beliefs about the importance or influence of

learning in these dimensions. Delivery was most strongly influenced by two external barriers – lack of time, and lack of resources. It was also moderately explained by a lack of personal knowledge and/or interest in the area. In each case, the barriers hindered delivery most strongly in the mental, emotional, and/or social health areas. They did not hinder delivery in the physical health area.

This suggests that health education teachers who perceive that they lack time and resources are most likely to teach physical health-related health education. It has been suggested that physical health delivery is favoured over delivery in other dimensions because health education teachers are more familiar, knowledgeable, comfortable, and/or interested in the physical health content (Sinkinson, 2011).

#### **4.6.1 Implications and future directions.**

The results of this study have general implications for the design and delivery of mental and emotional health education programming in the New Zealand school context. They suggest that a perceived lack of time and available resources is limiting delivery in these areas. They also suggest that approximately one-third of all primary and intermediate schools utilise external providers in their provision of health education classes. It is possible that these external providers are being hired in order to address concerns about time and resource shortages. The results of this study do not however, explore reasons for their use, nor do they investigate the impact that externally provided health programmes have on children's attitudes, behaviour, or the health outcomes.

From the social health perspective, these results suggest that children in primary schools are unlikely to receive regular, teacher-led instruction in social health education. The results also suggest that this may be due to a lack of time and resources, and that external providers of social health promotion programmes are likely to be used to deliver social health education. Yet little is known about the implementation or impact of these programmes. A two-stage evaluation study was therefore designed to investigate the use of one external provider of social health education programming in New Zealand. Kiwi Can, a programme that operates in low-decile primary and intermediate schools, was selected as the case study.

## **Chapter 5: The Implementation Evaluation of Kiwi Can**

### **5.1 Introduction**

This chapter is divided into two parts. In Part one (section 5.2), the ‘case’ under investigation in this PhD, Kiwi Can, is described. In Part two (section 5.3), an overview of the design and methods employed to evaluate the implementation of the Kiwi Can programme is provided. The chapter concludes with a discussion of the results of the implementation evaluation study, and the implications of those results.

### **5.2 The Kiwi Can Programme**

In this PhD, an instrumental case study approach was selected to test the impact that a fully resourced, time efficient, social health promotion programme had on the social attitudes and behaviours of primary and intermediate school children from low-decile communities around New Zealand. According to Punch (2009), instrumental case studies involve the in-depth study of a single case (i.e., a person, specific group, occupation, department, or organisation) “in order to provide insight into a particular issue, to build theory, or re-draw generalisations” (p. 121). In this thesis, an instrumental case study approach was selected to elicit insights into the impact that participation in a school-based health promotion programme could have on student social attitudes and behaviours. The single case under investigation was Kiwi Can, a positive youth development and health promotion programme. Kiwi Can was chosen for a number of reasons – it operates in low-decile primary and intermediate schools around New Zealand (and in many regions is also the main external provider of social health education programming), and it is delivered to all children aged five to 13 in participating schools on a weekly basis. In addition, the framework for the Kiwi Can programme was informed by the key positive youth development theory underpinning this thesis – Lerner’s 5C’s theory of positive youth development. From an implementation perspective, the Kiwi Can programme also had a lengthy and established history of working with schools, and was likely to continue operating for the duration of the PhD study. Finally, the Kiwi Can programme delivery team was actively seeking feedback on the provision and impact of their programme. For each of these reasons, it was expected that the Kiwi Can programme would be a suitable

case for use in evaluating the integrity of a positive youth development-oriented approach to social health promotion in New Zealand schools.

Kiwi Can is a values and life skills (or asset-based) programme, and is informed by positive developmental theory. At the time of publication, Kiwi Can had been operating in low-decile (Decile 1-3) primary and intermediate schools throughout New Zealand for twelve years. It was originally designed by a small not-for-profit Christian organisation in Auckland, before being subsumed by FYD in 2004. FYD is a charitable trust that was established in the mid-1990s by outdoor enthusiasts Graeme Dingle and Jo-anne Wilkinson (FYD, 2014). The original goal of the trust was to work with local communities to improve youth health and wellbeing through the provision of positive developmental programming. The Foundation oversees the delivery of five school-based programmes – Project K, MYND, Stars, Career Navigator, and Kiwi Can. Kiwi Can is the only one of the five to operate in both primary and intermediate schools. During the period of this PhD (2011-2014), it operated in schools in eight geographically and culturally diverse regions of New Zealand – Northland ( $n = 8$  schools), Auckland ( $n = 12$  schools), Waikato ( $n = 6$  schools), Coromandel ( $n = 2$  schools), Bay of Plenty ( $n = 14$  schools), Hawke's Bay ( $n = 2$  schools), South Canterbury ( $n = 8$  schools), and Southland ( $n = 12$  schools). In each of these regions, a licensed community partner was responsible for programme provision, including the hiring and training of staff (called *leaders*). Kiwi Can leaders were typically between 18 and 30 years of age, and resided in the local community. Most had between one and 11 years of experience as a Kiwi Can leader, and none had been formally trained as a teacher. To prepare for their role, leaders attended a compulsory week-long residential training programme, held annually at the Scouts Camp in Blockhouse Bay, Auckland. The training programme included workshops on lesson planning, programme delivery, behaviour management, resource design, teamwork, and communication. Two additional training sessions were also held each year in July and October. These were smaller in size and were offered in four regions around the country – Auckland, Waikato, Hawke's Bay, and Southland. Leaders from each of the eight regions were assigned to one of the four sessions and travelled to attend. Members of FYD's Curriculum and Pedagogy and Research and Evaluation teams facilitated both national and regional training. These teams were composed of experts in the fields of curriculum design and evaluation theory, who had, in the past,

been responsible for designing and evaluating the efficacy of participation in the Kiwi Can programme. They had also actively partnered with two external organisations, the MOE and The University of Auckland, to evaluate the context (theory) and implementation (delivery) of the programme.

The Kiwi Can programme has been externally evaluated on two occasions. The first evaluation was a review of the delivery of the programme. It was completed in 2002 by a research team commissioned by the MOE. The goals of the evaluation were to gather data on the sustainability, transferability, and adaptability of the Kiwi Can programme. Results indicated that the programme was reaching its target audience (students in low-decile schools), that Kiwi Can schools were meeting their implementation goals for the programme, and that school leaders believed that the programme was effective (Murrow, Kalafatelis, Fryer, Ryan, & Downen, 2002). Although the results were positive, they were largely anecdotal in nature, collected entirely from school staff (and not programme participants), and did not include quantitative measures of quality, fidelity, dosage, or participant responsiveness.

Given these omissions, a second, larger evaluation study was designed by a research team at The University of Auckland in 2008 (Ahmed, 2010). The study had three aims – (i) to conduct a context evaluation of the Kiwi Can programme (in keeping with the CIPP model of programme evaluation); (ii) to use the results to inform the design of a comprehensive implementation evaluation; and (iii) the design of a comprehensive outcome evaluation of the programme. Unfortunately, the research team was unable to complete either the implementation or outcome evaluations. They were, however, able to finish the context evaluation and produce a comprehensive programme logic model (Ahmed, 2010). A logic model is a schematic diagram that depicts the expected outcomes of participation in a particular programme over time, given the assumed inputs of the intervention. It is common for the outcomes to be divided into short, medium and long-term goals of participation in a given programme (Ahmed, 2010). According to the model, children involved in the Kiwi Can programme were expected to realise a range of positive youth development outcomes. In the short term, they were expected to improve their social skills, and become competent at regulating their own behaviour and emotions. In the medium term, they were expected to develop skills and attitudes that promoted healthy social, emotional,

mental, and spiritual development. In the long term, they were expected to realise each of Lerner's 5C's of positive youth development and become confident, competent, connected, caring individuals of high moral character who could both cope successfully with challenges in their lives and behave in health promoting ways. This logic model was adopted by the Curriculum and Pedagogy team at FYD and used to re-design the Kiwi Can curriculum. A new programme curriculum was completed in 2011. See Appendix F for a copy of the current Kiwi Can curriculum.

As of 2011, the Kiwi Can programme consisted of both skill-based and knowledge-based lessons. The lessons were delivered to all children at participating low-decile schools on a class-by-class basis, in a specially designated classroom, located on school grounds. Kiwi Can classes varied in length from 30 minutes for New Entrants, to 60 minutes for Year 7 and Year 8 students, and each class was facilitated by two Kiwi Can leaders (usually a male and a female). All the children in the school participated in one Kiwi Can lesson per week for the duration of the academic year (i.e., 40 lessons per year). Most participating schools had had Kiwi Can on site for many years, but other schools had only recently adopted the programme. Years of experience with the programme varied both between and within regions.

Kiwi Can had a national curriculum document, underpinned by four specific principles, and designed around four specific values. The values – Positive Relationships, Integrity, Resilience, and Respect, served as key themes, and were comprised of six modules, that were then delivered as units of study. Curriculum content and delivery was standardised throughout the country. The units were designed to be four weeks in length, and it was expected that schools would cover two themes and twelve modules in a single academic year. Schools, therefore, had to participate in the programme for two full academic years before children would have received instruction in every area. Schools were therefore considered to be *experienced* when they had participated in the Kiwi Can programme for more than two years. Prior to this, they were labelled *new* schools. At the end of the two-year period, schools returned to the first theme and repeated the four-theme cycle. This meant that children who attended an experienced Kiwi Can school from Year 1 to Year 8 were exposed to learning in each thematic area four times over the course of their lives. See Table 16 for a list of the themes ( $n = 4$ ) and modules ( $n = 24$ )

contained in the current Kiwi Can curriculum. All four themes are based on values that promote prosocial attitudes and behaviours.

Table 16

*Themes and Modules Present in the Kiwi Can Programme Curriculum*

Themes	Positive Relationships	Integrity	Resilience	Respect
Modules	Positive Communication Cooperation	Honesty	Understanding Emotions	For Ourselves
	Friendship	Responsibility	Self-control /Discipline	Respectful Communication
	Fairness and Fair Play	Reliability	Dealing With challenges	For Others
	Leadership	Making Good Choices	Goal Setting	For Our School
	Conflict Resolution	Being a Role Model	Problem Solving	For Our Community
		Accountability	Perseverance	For Our Environment

*Note.* Table is adapted from the “Kiwi Can Co-ordinator Manual” by The Foundation for Youth Development, 2012, p. 3.

Through repeated exposure to the Kiwi Can programme, FYD posited that children would develop: (i) a “can do” attitude to the challenges of life, and a positive self-belief in their capabilities; (ii) a sense of self-worth, and greater self-confidence; (iii) respect for themselves, each other, their school, family, local community, and New Zealand; (iv) a set of core values, and the ability to take responsibility and be accountable for their actions; and (v) the skills to make positive, healthy decisions about their behaviour (Ahmed, 2010). The research team at FYD however, had yet to empirically test for programme outcomes, or to evaluate the quality, fidelity, and/or participant responsiveness to programme delivery, in any of the eight Kiwi Can regions. A two-stage, sequential programme evaluation of Kiwi Can was therefore designed. In Stage 1, the *quality, fidelity, and participant responsiveness* associated with the delivery of the programme was evaluated. These results informed the design of an outcome evaluation that in Stage 2, was used to investigate the impact that participation in Kiwi Can had on children’s social attitudes and behaviours. The design, methods, and results of the first of these two evaluations, the implementation evaluation, are presented in the second half of this chapter. The design, methods, and results of the outcome evaluation are presented in the chapter that follows.

### **5.3 The Implementation Evaluation of Kiwi Can – Introduction**

The first stage of the evaluation of the Kiwi Can programme involved an investigation into programme delivery (or implementation). The purpose of this study was to ascertain the quality, fidelity, and degree of participant responsiveness associated with the delivery of the programme in each of the eight participating Kiwi Can regions. These three criteria were drawn directly from the work of Durlak and DuPre (2008), who loosely defined fidelity as being a measure of how well programme delivery adheres to the originally intended programme, quality as being how well programme components are delivered, and participant responsiveness as being how stimulating the programme is and how well it sustains interest. It was assumed that content and delivery of the Kiwi Can programme would be the same in all participating schools, and that quality, fidelity, and participant responsiveness would not differ according to the sex, year level, or ethnicity of participants, or to school condition (in this study school condition referred to a school's status as new or experienced – schools were deemed to be new if they had less than two full years of experience with the Kiwi Can programme, and experienced, if they had more than two).

The study was guided by three main research questions –

1. According to Year 4-6 student participants, how well was the Kiwi Can programme implemented in low-decile primary schools across New Zealand? (Quality)

1(i) To what degree did Year 4-6 student participants report learning new skills and values? (Quality)

1(ii) To what degree did Year 4-6 student participants report satisfaction with the Kiwi Can programme structure, organisation, and leadership? (Quality)

1(iii) To what degree did Year 4-6 student participants report enjoying the Kiwi Can programme? (Participant responsiveness)

2. How frequently did students report learning about prosocial values and practicing social skills in their Kiwi Can classroom during the 2012 school year? (Fidelity)

3. Did participant responses differ by school condition, year level, sex, and/or ethnicity?

## 5.4 Study Design

The implementation evaluation of Kiwi Can employed a non-experimental, cross-sectional design. Data were collected using an anonymous survey. Children in Year 4 to Year 6 from new and experienced Kiwi Can schools across New Zealand were selected for participation according to a purposive, cluster sampling technique. This technique was chosen for logistical and financial reasons (there was a tight timeline and small operating budget for the project). Children in Years 4 to 6 were targeted because they were deemed sufficiently literate to be able to complete the survey on their own, and because all Kiwi Can schools had children in Years 4 to 6.

### 5.4.1 Sampling frame.

The sampling framework consisted of eight schools (one each from the eight participating Kiwi Can regions) distributed as three from Year 4, two from Year 5, and three from Year 6 (see Table 17). Initially the eight regions were randomly divided into new and experienced categories. Schools within each region were also categorised as new or experienced depending on the number of years that they had participated in the programme.

Table 17

*Sampling Frame for the Kiwi Can Implementation Evaluation*

Year Group	Kiwi Can Status		Number of Groups
	Experienced	New	
4	Northland; Southland	Bay of Plenty	3
5	Hawke's Bay	Waikato	2
6	Coromandel	Auckland; South Canterbury	3
Total	4	4	8

On the basis of this framework, a list of potential schools ( $n = 52$ ) was drawn from the national Kiwi Can database. Of those schools, 33 were eligible for inclusion in the sampling frame (13 experienced and 20 new schools). The names of each school were then written on pieces of paper and arranged into groups by region. One piece of paper (i.e., one school) was then randomly selected from the pile for each region.

This produced a final list of eight participating Kiwi Can schools, one from each region.

All Year 4, 5, and 6 children from the selected schools were invited to participate in the study by their Kiwi Can leaders. Permission to participate was sought from children and from their parents. All the children were granted permission by their parents to participate in the anonymous survey study. See Appendices G and H for copies of the Student Information Sheet and Parental Non-Consent form for this study. Data collection commenced in late October 2012, and concluded in late November 2012. During this time, all participants completed a 35-item questionnaire, which took about 15 minutes to complete.

#### 5.4.2 Participants.

A total of 319 children participated in the anonymous survey study (see Table 18), and numbers from new and experienced schools were roughly equal. There were slightly more female than male participants, and more children from Year 4 (35.4%), than from Year 3 (10.3%), Year 5 (25.7%), and Year 6 (28.5%). Year 3 participants were included as they attended Kiwi Can as part of integrated Year 3-4 classes. Just over half of the participants identified as Māori, and a third as New Zealand European.

Table 18

*Frequency Count of Participants by Demographic Variable*

Variable	Frequency ( <i>n</i> )	Percentage (%)
School Condition		
Experienced	166	52.0
New	153	48.0
Sex		
Male	154	48.3
Female	165	51.7
Year Level		
Year 3	33	10.3
Year 4	114	35.4
Year 5	82	25.7
Year 6	91	28.5
Ethnicity		
Māori	173	54.2
Pakeha/NZ European	103	32.3
Pasifika	24	7.5
Asian	8	2.5
Other	11	3.4

### 5.4.3 Instrument.

A 35-item questionnaire was used to gather data for the implementation evaluation and its design was informed by the literature, informal observations of the Kiwi Can programme, and discussions with key stakeholders (i.e., Kiwi Can schools, and employees of FYD). The questionnaire was piloted with five children from Year 4 to Year 6 from a school in South Auckland, to ensure clarity of instructions and content. No issues were reported. The pilot data were not included in the main research project.

The final Kiwi Can implementation evaluation questionnaire consisted of three parts. Part A contained five demographic questions (sex, age, year level, number of years of schooling, and ethnicity). For ethnicity, students were permitted to choose from a list of five options– Māori, Pasifika, Pakeha/New Zealand European, Asian, and Other. For analysis purposes, the standard Statistics New Zealand prioritisation method was used to assign students to a single ethnicity category (Lang, 2002; Statistics New Zealand, 2005). This method is based on a formula: Māori < 2.Pasifika < 3.Asian < 4.Pakeha < 5.Other. Children who indicated that they were Māori (and any other ethnicity) were assigned to the Māori category. Children who indicated that they were Pasifika (and any other category except for Māori) were assigned to the Pasifika category. Children who indicated that they were Asian (and any other category except for Māori or Pasifika) were assigned to the Asian category. Lastly, children who shaded the Pakeha/New Zealand European box (and Other, but not Māori, Pasifika, or Asian) were assigned to the Pakeha category.

In Part B, children were then asked to re-call whether they had learned eight prosocial values and practiced six prosocial skills in Kiwi Can lessons. The eight values included were – respect, honesty, responsibility, motivation, capacity to cope with challenges, politeness, being a good friend, and being a good team member. The six social skills were – making goals, solving problems, talking nicely to other people, being kind to others, co-operating with others, and taking turns.

In Part C, participants were asked to rate the degree to which they agreed with 16 statements regarding participation in the Kiwi Can programme, using a 5-point rating scale (see Table 19). The statements were grouped into five latent factors of between

two and four items – Liking Kiwi Can ( $n = 4$  items), Learning in Kiwi Can ( $n = 3$  items), Lesson Structure ( $n = 4$  items), Resources ( $n = 3$  items), and Kiwi Can Leaders ( $n = 2$  items). The item-response options were: 1 = *strongly disagree*, 2 = *disagree*, 3 = *slightly agree*, 4 = *mostly agree*, and 5 = *strongly agree*. See Appendix I for a copy of the Kiwi Can implementation evaluation questionnaire.

Table 19

*Latent Factors and Items from the Kiwi Can Implementation Evaluation*

*Questionnaire*

Factor and Item
Liking Kiwi Can
Overall, I really like the Kiwi Can lessons
I am often bored during Kiwi Can lessons
I participate in all aspects of the Kiwi Can lesson
Kiwi Can lessons are the best part of the school week
Learning in Kiwi Can
I have learned new things from taking part in the Kiwi Can lessons
Kiwi Can lessons teach me important skills
Kiwi Can lessons teach me about values
Kiwi Can Lesson Structure
The Kiwi Can lessons follow the same pattern (Recap, Energiser, Catch Phrase and Activity)
We have a topic in each Kiwi Can lesson
My Kiwi Can lessons start on time
My Kiwi Can lessons are well organised
Resources
We use equipment in our Kiwi Can lessons
There is enough equipment for everyone in the class
I am taught how to use the Kiwi Can equipment safely
Kiwi Can leaders
I look up to my Kiwi Can leaders
My Kiwi Can leaders are energetic

## 5.5 Data Preparation and Analysis

Manual inspection of all questionnaires for mono-numeric, patterned responding revealed no such patterns. Missing values analysis indicated that the data were missing at random (Little's MCAR  $\chi^2 = 728.372$ ,  $df = 507$ ,  $p = .11$ ), therefore, missing values were imputed using EM. The final data set consisted of 319 complete student questionnaires, representing approximately 1.0% of all children currently enrolled in Kiwi Can from new entrant level to Year 8, and 14.5% of all children enrolled in Year 4 to Year 6.

Data from Part B was analysed using a k-means cluster analysis technique. Pearson's two-sided chi-square tests and odd ratios were generated to describe differences in

frequency between children as a function of school condition, year level, sex, region, and ethnicity. CFA was used to test the proposed factor model for Part C. Once scales were established, significant mean differences in factor endorsement was evaluated for year level, sex, ethnicity, region, and school condition using MANOVA. The magnitude and direction of effect was determined using the Pillai-Bartlett trace and Games-Howell post hoc test statistics.

## **5.6 Results - Recalled Values and Practices (Part B)**

The iterative, k-means analysis of the dichotomous (yes/no) categorical selection of items in Part B supported the initial hypothesis that there were, indeed, two distinct clusters of data in Part B (Cluster 1 - *Learning Values*, ( $n = 252$  cases), and Cluster 2 - *Social Skill Development* ( $n = 63$  cases). The iteration score for Cluster 1 was 1.38, and the iteration score for Cluster 2 was 1.15, based on a five-stage iterative process. On this basis the data were then analysed and reported as two separate clusters – Learning Values (Cluster 1) and Social Skill Development (Cluster 2).

Overall, the students recalled being taught or practicing both the Learning Values and Social Skills Development aspects, with positive response rates averaging 80% or better. There was a moderate difference ( $d = .51$ ) in favour of Social Skills Development over the Learning Values. The most anomalous result was the very low level of recall for the learning value, motivation (see Table 20).

Table 20

*Frequency of Student Responses for the Learning Values and Social Skill Development Data from the Kiwi Can Programme Questionnaire (Part B)*

Values and Practices	Frequency - Yes	Frequency (%)
<b>Learning Values</b>		
Honesty	297	93.1
Respect	309	96.9
Responsibility	261	81.8
Motivation	156	48.9
Coping with challenges	233	73.0
Being polite	242	75.9
Being a good friend	276	86.5
Being a good team member	279	87.5
<b>Average</b>	256.63 ( <i>SD</i> = 48.06)	80.45 ( <i>SD</i> = 15.07)
<b>Social Skills Development</b>		
Making goals	265	83.1
Solving problems	272	85.3
Talking nicely to others	269	84.3
Being kind to others	286	89.7
Co-operating with others	278	87.1
Taking turns	274	85.9
<b>Average</b>	274 ( <i>SD</i> = 7.35)	85.90 ( <i>SD</i> = 2.31)

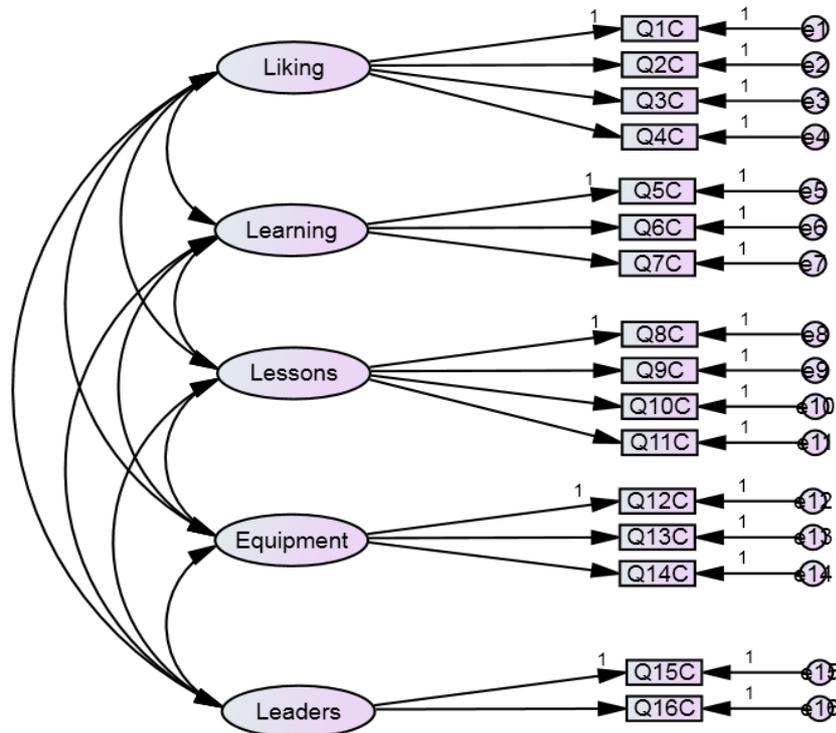
The results of the Pearson's two-sided chi-square tests indicated that there were only three statistically significant differences in these values and practices. Only school condition had a statistically significant effect for responsibility and motivation.

Experienced school students reported learning about being responsible more frequently than new school students,  $\chi^2_{(1)} = 10.55, p < .001, OR = 2.71$  and they also remembered practicing solving problems more than new school students,  $\chi^2_{(1)} = 6.14, p < .02, OR = 2.26$ . In contrast, new school students reported learning to be motivated more frequently than experienced school students,  $\chi^2_{(1)} = 11.00, p < .001, OR = 2.12$ . Otherwise, there were no statistically significant associations for all other learning areas or social skills by school condition, sex, year level, or ethnicity.

### 5.7 Results - Kiwi Can Participation (Part C)

The validity of the intended five-factor model of Kiwi Can participation in Part C was tested with CFA (see Figure 1). The model tested was a simple inter-correlation of the five factors, each predicting the items intended for it, as shown in Table 19. The underlying covariance matrix was not positive definite, and inadmissible. This suggested that too many factors were present, and so further simplification of dimensionality was carried out.

**Figure 1.** The Original Five-Factor Structural Model



*Figure 1.* A copy of the original measurement model used to test the validity of the original, five-factor structure in Part C of the Kiwi Can implementation evaluation.

It was decided to merge the three programme factors (i.e., Kiwi Can Lesson Structure, Resources, and Kiwi Can Leaders) into a single latent factor (Kiwi Can Programme) because they all referred to general characteristics of the programme. A new three-factor model (see Figure 2) was tested in CFA. This model had acceptable to good fit characteristics ( $\chi^2/df = 2.47$ , CFI = .90; RMSEA = .068, 90% CI [.057, .079], SRMR = .049; gamma hat = .94). The three factors were moderately to highly correlated ( $r = .76 - .86$ ), and standardised item loadings ranged from moderate ( $\beta = .37$ ) to high ( $\beta = .80$ ).

**Figure 2.** The Final Three-Factor Model

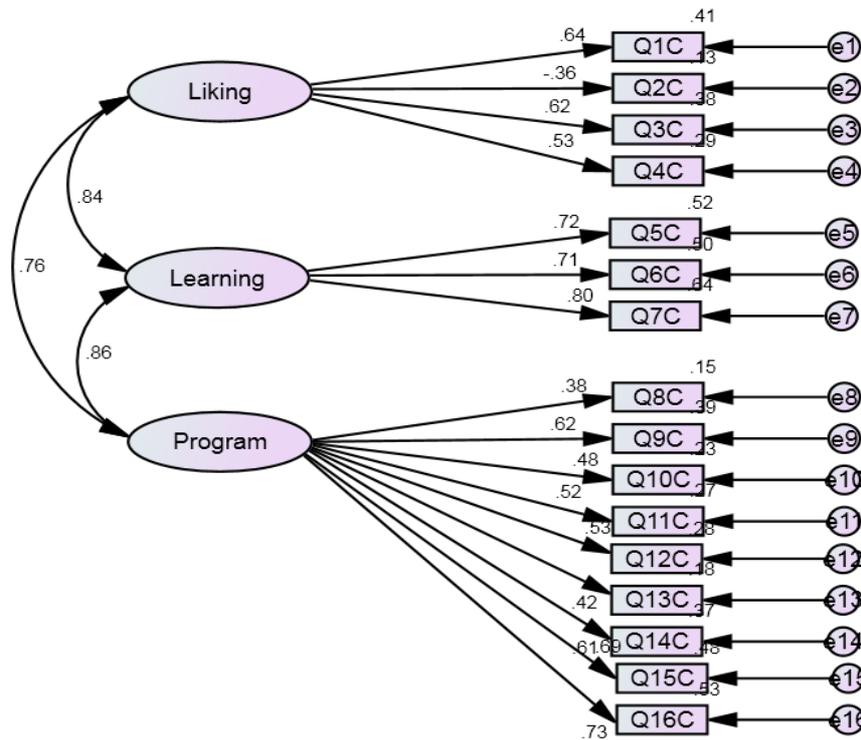


Figure 2. A schematic diagram of the final three-factor structural model for part C of the Kiwi Can implementation evaluation student survey. Standardised values shown on regression paths and correlations; proportion variance explained shown on error term paths.

The new model had measurement, metric, and scalar invariance across sex ( $\Delta\text{CFI} < .01$ ) and school condition ( $\Delta\text{CFI} < .01$ ) (see Table 21). It was not possible to test invariance across ethnicities or year levels given the small sample sizes in some conditions ( $n \leq 50$ ).

Table 21

*Invariance Testing for Three-Factor Structural Equation Model by Condition*

Level of Constraint - SEM	CFI (Sex)	$\Delta\text{CFI}$ (Sex)	CFI (School)	$\Delta\text{CFI}$ (School)
Unconstrained model	.832	—	.828	—
Measurement weights	.826	.006	.821	.007
Measurement intercepts	.822	.004	.818	.003
Structural covariances	.817	.005	.809	.009
Measurement residuals	.815	.002	.799	.010

This analysis established three latent factors concerning participation in Kiwi Can that were used for subsequent analysis. Overall, self-reported mean scores were highest for Learning in Kiwi Can, followed by Kiwi Can Programme, and finally Liking Kiwi

Can (see Table 22). However, the difference between Learning in Kiwi Can and the Kiwi Can Programme was trivial ( $d = .16$ ), and very large for both of these compared to Liking Kiwi Can ( $d = 1.56$  and  $d = 1.60$ , respectively). When mean differences were explored, the Pillai-Bartlett trace indicated that there were no statistically significant differences in the degree to which Year 4-6 children Learned in Kiwi Can, Liked Kiwi Can, or viewed the characteristics of the Kiwi Can programme as a function of age,  $V = .08$ ,  $F_{(12, 942)} = 2.15$ ,  $p = .12$ , year level,  $V = .03$ ,  $F_{(9, 945)} = 1.18$ ,  $p = .31$ , sex  $V = .02$ ,  $F_{(3, 315)} = 2.22$ ,  $p = .09$ , or ethnicity  $V = .03$ ,  $F_{(12, 942)} = .77$ ,  $p = .69$ . There was a small but significant difference, however, by school condition. In this case new children reported lower mean scores for Kiwi Can programme characteristics than experienced children,  $V = .04$ ,  $F_{(3, 315)} = 4.52$ ,  $p < .01$ ,  $d = .21$ .

Table 22

*Mean Scores for Liking Kiwi Can, Learning in Kiwi Can, and Kiwi Can Programme by School Condition, Sex, Year Level, and Ethnicity*

Variable	<i>n</i>	<u>Liking KC</u> <i>M (SD)</i>	<u>Learning in KC</u> <i>M (SD)</i>	<u>KC Programme</u> <i>M (SD)</i>
Total	319	3.71 (.50)	4.65 (.69)	4.55 (.55)
School Condition				
New KC	164	3.73 (.45)	4.68 (.66)	4.49 (.65)
Experienced KC	155	3.68 (.55)	4.62 (.72)	4.62 (.42)
Sex				
Male	154	3.66 (.56)	4.57 (.76)	4.47 (.60)
Female	165	3.75 (.44)	4.65 (.69)	4.55 (.55)
Year Level				
Year 3	33	3.61 (.71)	4.61 (.89)	4.68 (.39)
Year 4	113	3.72 (.48)	4.66 (.69)	4.58 (.43)
Year 5	82	3.68 (.39)	4.69 (.53)	4.58 (.46)
Year 6	91	3.73 (.54)	4.61 (.77)	4.45 (.77)
Ethnicity				
Māori	173	3.70 (.55)	4.65 (.73)	4.52 (.63)
Pasifika	24	3.82 (.39)	4.69 (.67)	4.52 (.58)
Pakeha	103	3.71 (.43)	4.65 (.64)	4.59 (.44)
Asian	8	3.75 (.27)	4.79 (.35)	4.54 (.45)
Other	11	3.57 (.62)	4.48 (.82)	4.70 (.20)

*Note.* Mean values based on a positively packed 5-point rating scale (1 = *disagree*, 2 = *slightly disagree*, 3 = *slightly agree*, 4 = *mostly agree*, 5 = *agree*). KC = Kiwi Can.

## 5.8 Conclusion

In Part C, participants reported that the Kiwi Can programme was delivered to a high standard, and this opinion was shared across the eight sampled schools. In fact, high mean scores ( $M > 4.00$ ) were given for both Learning in Kiwi Can and for Kiwi Can Programme Characteristics. This suggested that the children believed they were learning a great deal in Kiwi Can, had adequate access to resources, and were supported by strong programme leaders. It also suggested that they felt that the Kiwi Can lessons were well organised, topical, and structured to meet their needs. There were small differences in mean scores depending on whether the school was new or experienced, but the size of the differences were small and only specific to the programme characteristics factor, suggesting that children were aware of school and/or leader inexperience with the programme. However, the small size of difference ( $d = .21$ ) suggested that the differences would likely disappear with time and practice.

In Part C, participants also reported substantially lower mean scores for liking, a measure of participant responsiveness ( $M = 3.71$ ,  $SD = .50$ ), than for learning ( $M = 4.65$ ,  $SD = .69$ ), or for Kiwi Can programme characteristics ( $M = 4.55$ ,  $SD = .55$ ). The lower scores for liking could be viewed as problematic given that some research has indicated that enjoyment is a prerequisite for learning and achievement (Abu-Hilal, 1992; Marsh, Hau, Artelt, Baumert, & Peschar, 2006; O'Brien, Martinez-Pons, & Kopala, 1999; Walker, 2003), though the finding is not universal. In fact, in a recent study on achievement amongst immigrants in New Zealand, Otunuku and Brown (2007) reported that for Pasifika and Maori children, there was a near zero correlation between liking and achievement (learning) in mathematics, writing, and reading. Further, Winnheller, Hattie, and Brown (2013) also reported that liking was not predictive of academic performance in reading, writing, or mathematics, amongst two samples of Year 8 ( $n = 336$ ) and Year 10 ( $n = 272$ ) children in New Zealand public schools. Whether 'liking' is problematic or not in the Kiwi Can context cannot be answered with the current data. It may be that the learning benefits outweigh the lack of interest, but further work is needed to determine whether a response by Kiwi Can is required.

The very high mean scores that children reported for learning in Kiwi Can and Kiwi Can programme characteristics also demand careful consideration. It is possible, given the age of the participants, that the high mean scores were merely a product of bias in reporting, particularly social desirability bias. Social desirability is a general tendency to overestimate desirable traits and behaviours when using self-report measures (Dadds, Peerin, & Yule, 1998). It is common in children (Reynolds & Paget, 1983) and in studies of health-related behaviours and attitudes (Klesges et al., 2004). It is also commonly reported in surveys of participant satisfaction with health programmes and initiatives (Mazor, Clauser, Field, Yood, & Gruwitz, 2002). Given that Kiwi Can is a health promotion initiative for children, it was possible that they may have succumbed to a social desirability bias and overrated the quality and learning that they experienced in the programme. The lower liking scores provided by participants however, suggest that the Kiwi Can students did not succumb to this bias, and that the very high mean scores for learning in Kiwi Can and Kiwi Can programme characteristics are likely to be valid, and not adversely affected by a drive to please their Kiwi Can leaders by reporting inflated mean scores.

The results from Part B, on the other hand, provide support for the notion that the Kiwi Can programme was sufficiently powerful to enable children to recall being taught or having practiced the aspects of the programme that were part of its explicit curriculum. The high average rates of recall for the prosocial values areas ( $M = 80.45\%$ ,  $SD = 15.07$ ) and social skills areas ( $M = 85.90\%$ ,  $SD = 2.31$ ) suggest that the Kiwi Can curriculum had high fidelity. In other words, children experienced lessons that adhered to the main programme themes as outlined in the curriculum. Given that 'motivation' was not a main programme theme (see Appendix D), it was not surprising that this was less frequently recalled than all other prosocial values. Further, given the lack of statistically significant differences by year level, sex, or ethnicity, there is evidence that the Kiwi Can programme curriculum was being consistently delivered across all sample schools.

Overall, the results of this survey suggest that the Kiwi Can programme was being consistently delivered to a high standard, with fidelity, and a focus on participant responsiveness across all sampled schools. The findings also suggest that children felt that they were learning about and practicing skills that promoted prosocial

attitudes and behaviours. It could therefore be concluded that the Kiwi Can programme was being well delivered, at least from children's perspectives.

This study did not however establish that the Kiwi Can programme actually helped to deliver more positive outcomes for its participants, and this is the focus of the outcome evaluation of Kiwi Can.

## Chapter 6: The Outcome Evaluation of Kiwi Can

### 6.1 Introduction

In this chapter the results of a year-long outcome evaluation of the Kiwi Can programme are presented. The overarching purpose of this evaluation was to assess the impact that participation in an externally provided, social health promotion programme had on the self-reported social attitudes and behaviours of children in low-decile schools. Changes in self-rated positive social development (PYD) and classroom climate (CC) dimensions were examined over the course of one academic year (2013). Measures of PYD and CC were also collected at baseline, and scores were analysed as a function of school condition, age, sex, ethnicity, and transience.

The study was guided by six main research questions –

1. What impact does participation in Kiwi Can have on the self-reported incidence of individual prosocial attitudes and behaviours, amongst vulnerable students in low-decile primary and intermediate schools?

1(i) Are there statistically significant differences in the measures of individual prosocial attitudes and behaviours as a function of school condition (i.e., new Kiwi Can, experienced Kiwi Can, and/or control school), sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools at baseline?

1(ii) Are there statistically significant differences in measures of individual prosocial attitudes and behaviours as a function of school condition, sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools over the course of one academic year?

2. What impact does participation in Kiwi Can have on the self-reported incidence of prosocial attitudes and behaviours at the classroom level, amongst vulnerable students in low-decile primary and intermediate schools?

2(i) Are there statistically significant differences in the measures of classroom climate as a function of school condition, sex, year level, and/or ethnicity

amongst vulnerable children in low-decile primary and intermediate schools at baseline?

2(ii) Are there statistically significant differences in the measures of classroom climate as a function of school condition, sex, year level, and/or ethnicity amongst vulnerable children in low-decile primary and intermediate schools over the course of one academic year?

3. What impact does school-level transience have on self-reported incidence of prosocial attitudes and behaviours at baseline and across time for children in low-decile primary and intermediate schools across New Zealand?

4. What impact does school-level transience have on self-reported measures of classroom climate at baseline and across time for children in low-decile primary and intermediate schools across New Zealand?

5. Are there statistically significant differences in measures of individual prosocial attitudes and behaviours or classroom climate for children as a function of school-level transience and school condition?

6. What is the magnitude and direction of the interaction between school condition and transience level on self-reported incidence of individual pro-social attitudes and behaviours, and measures of classroom climate for children at baseline and across time?

According to the literature, children engaged in positive youth development programmes like Kiwi Can are expected to develop health promoting attitudes and values, and assets like interpersonal and conflict resolution skills. As these values and skills develop, children are then expected to become more caring, compassionate, connected, and socially competent, and to experience prosocial development and positive social health outcomes. This study tested that theory in a sample of children from low-decile primary and intermediate schools in New Zealand.

## 6.2 Methods - Strategy and Design

In this study the Kiwi Can programme served as a case setting for investigating the impact that a positive youth development and health promotion programme had on the social attitudes and behaviours of children from low-decile primary and intermediate schools in New Zealand. A quasi-experimental, pre-post survey design was used. Participants were either from schools participating in Kiwi Can, constituting the treatment, or in schools not affiliated with Kiwi Can, which functioned as *control* for the effect of the Kiwi Can treatment. Schools involved in the Kiwi Can programme for more than two years pre-evaluation were classified as experienced Kiwi Can schools. Those involved in the programme for less than two years were classified as new Kiwi Can schools. Therefore, for this study there were three groups in total – experienced Kiwi Can, new Kiwi Can and control schools.

This design enabled the testing of four hypotheses –

1. Experienced Kiwi Can students will report higher mean scores for both positive social development in youth (PYD) and prosocial classroom climate (CC) outcomes (subscales), than either new Kiwi Can or control school students at baseline.
2. Students from experienced Kiwi Can and new Kiwi Can schools will report gains in mean PYD and CC subscale scores over the course of the academic year, and new Kiwi Can school students will report the largest mean increase in scores for both constructs over time.
3. Pre-test and post-test mean scores for PYD and CC will not be statistically significantly different for control school students.
4. Because Kiwi Can seeks to treat all students equally regardless of sex, age, or ethnicity, pre-test and post-test mean scores will be statistically equivalent across these demographic characteristics.

### 6.2.1 Instrument.

The data for this study were collected using a self-report questionnaire (see Appendix J). In Part A of the questionnaire students were asked to respond to six demographic

questions related to their age, sex, year of schooling, years in current school, ethnicity, and participation in the Kiwi Can programme.

In Part B, 40 questions on student social attitudes and behaviours were posed (see Table 23). Social development was inferred from four socially-orientated scales adopted from Lerner's 5C's model of positive youth development (i.e., Character, (social) Competence, Caring/Compassion, and Connection). Each scale was subdivided into two or three subscales consisting of three to five items. The items for Character, Caring/Compassion, and Connection were taken from the PSL: AB (Search Institute, 2012), which has been used to assess positive youth development outcomes in American youth. It was selected because of previous evidence of construct validity and scale reliability for the Caring/Compassion, Character, and Connection subscales (Bowers et al., 2010; Phelps et al., 2009; Theokas et al., 2005). Items for (social) Competence were taken from the three subscales (i.e., Social Rules, Likeability, and Social-Ingenuousness) in the Children's Self-Reported Social Skills Scale (CS<sup>4</sup>), a 21-item test of self-reported social skills (Danielson & Phelps, 2003). These scales had high estimates of reliability (i.e., Cronbach  $\alpha$  = .90, .86, and .84 respectively), in contrast to the low reliability (Cronbach's  $\alpha$  = .59) of the (social) Competence subscale in the PSL: AB (Jelicic et al., 2007). The remaining nine items from the original CS<sup>4</sup> tool were not selected for this study because they did not pertain to social development. All items in Part B used a positively packed 6-point rating scale (i.e., 1 = *Strongly disagree*, 2 = *Mostly disagree*, 3 = *Slightly disagree*, 4 = *Slightly agree*, 5 = *Mostly agree*, and 6 = *Strongly agree*).

Table 23

*Scales, Subscales, and Sample Items for the Positive Social Development, PYD, Section of the Kiwi Can Outcome Evaluation Questionnaire*

Subscale	Source	Number of items	Sample item
<b>Caring/Compassion</b>			
Empathic Concern	PSL: AB	5	I often feel sorry for other children who are sad or are in trouble
Perspective Taking	PSL: AB	5	I believe that there are two sides to every question and try to look at them both
<b>Social Competence</b>			
Social Rules/Politeness	CS <sup>4</sup>	5	I say thank you when someone does something nice for me
Likeability	CS <sup>4</sup>	3	Others like me and have fun with me
Social-ingenuousness	CS <sup>4</sup>	3	I kick or hit someone else when they make me angry
<b>Connection</b>			
Connection to Peers	PSL: AB	5	I trust my classmates
Connection to School	PSL: AB	5	I get along well with my teachers
<b>Character</b>			
Values Diversity	PSL: AB	4	I respect the values and beliefs of people who are of a different ethnic background/culture than I am
Personal Values	PSL: AB	5	I stand up for what I believe in, even when it's unpopular to do so

*Note.* PSL: AB = Profile of Student Life: Attitudes and Behaviour Survey; CS<sup>4</sup> = Children's Self-Reported Social Skills Scale.

In Part C, children were asked to respond to 22 statements about the behavior and attitudes of other children in their classroom in order to gather information about classroom climate (see Table 24). These items were drawn from the seven factors of the Measure of School Social Climate tool used to assess classroom climate in Canadian primary schools following the introduction of a bullying prevention programme (Ding et al., 2011). Five of the seven factors were selected – Positive Behavior in the Classroom, PB ( $n = 10$  items), Negative Behaviour in the Classroom, NB ( $n = 5$  items), Classroom and School Supportiveness, CSS ( $n = 5$  items), Safety at School, SS ( $n = 2$  items), and Enjoyment of School, ES ( $n = 4$  items). Two scales were not used because they asked questions about the behavior of the teacher, which this study did not seek to address. One item in the Enjoyment of School scale (“I feel safe on the playground and on the school grounds”) was separated into two separate items (“I feel safe on the playground” and “I feel safe on the school grounds”) to ensure that each factor had three or more items.

Table 24

*Scales, Subscales, and Sample Items for the Classroom Climate, CC, Section of the Kiwi Can Outcome Evaluation Questionnaire*

Scale	Number of items	Sample item
Positive Behaviour in the Classroom	5	Students in my class share with other students
Negative Behaviour in the Classroom	5	Students in my class borrow things without asking
Classroom and School Supportiveness	5	Students in my class work together to solve problems
Safety in the School	3	Students in my school feel safe in all areas of the school building
Enjoyment at School	4	I like my school

### **6.2.2 Sampling.**

All schools were selected using stratified, purposive sampling. The sampling frame incorporated two Kiwi Can schools (where possible) and one control school from each of the eight participating Kiwi Can regions. The regions were also classified as either new Kiwi Can or experienced Kiwi Can regions. The decision to select only two schools per region was made because the demographic characteristics of children tend to vary more widely within than between schools in New Zealand (Ditch & Tarleton, 2010). Only three year groups (Years 6, 7, and 8) were included in the survey because it was assumed that these children would be sufficiently literate to comprehend and complete the survey on their own. Additionally, previous Kiwi Can evaluations had focused on predominantly younger children, and this was an opportunity to examine the impact of the programme on older children.

A total of 64 schools were participating in the Kiwi Can programme at the beginning of 2013, and of those schools, 48 were included in the sampling frame – 34 experienced, and 14 new schools. Two new or two experienced schools from each of the eight Kiwi Can regions were then randomly selected according to the frame. It must be noted that four regions (Hawke's Bay, Coromandel, South Canterbury, and Waikato) had only one or two schools that fell into the new or experienced category, and so these schools were automatically selected for participation. Also, regions with new schools were automatically assigned to the new region category, and in three regions (Northland, Coromandel, and Southland), one or more of the randomly

selected schools had also participated previously in the Kiwi Can implementation evaluation.

The control schools were selected using a similar method to the Kiwi Can schools. One control school per Kiwi Can region was randomly selected according to a stratified sampling frame. That frame included all mainstream, co-educational, Decile 1-4, full primary, composite, and intermediate schools located within a 10 kilometre radius of a participating Kiwi Can school in each of the eight Kiwi Can regions. The TKI (Ministry of Education) website was used to generate the list of potential schools. One school was then randomly selected per region. When random selection of schools was complete, the total number of Year 6 to Year 8 children was estimated, and second and third schools in each region were then randomly selected to increase the number of control school children. The sampling framework can be found in Table 25.

Table 25

*Sampling Framework for Kiwi Can Outcome Evaluation*

School	R1	R2	R3	R4	R5	R6	R7	R8	Total
New				XX	XX	X		XX	7
Experienced	XX	XX	XX				XX		8
Control	WD	XX	X	X	X	XX	X	X	9
Total	2	4	3	3	3	3	3	3	24

*Note.* R = region, WD = withdrawn, X = a school, XX = two schools.

Kiwi Can and control schools agreed to participate on the condition of anonymity, and therefore no list of participating schools is provided in this thesis.

### **6.2.3 Recruitment procedures.**

*Kiwi Can schools.* Once the Kiwi Can schools had been selected, Kiwi Can co-ordinators in each region were asked to meet with principals to inform them about the study and request permission to conduct the research. All eight of the principals agreed, and signed the requisite paperwork. With principal consent, the Kiwi Can leaders began the process of recruiting children. Children were informed about the research and Parent and Student Participant Information sheets (PIS) as well as Parental Non-Consent forms (NCF) were distributed (see Appendices K - N for copies of these documents along with a sample PIS for school principals). The children were asked to take the paperwork home and return the NCF only if their parent or guardian

did not want them to participate. These children would be given an alternative task to complete during the time scheduled to complete the survey (see Appendix O for a copy of the alternative task). No Parental NCFs were returned.

*Control Schools.* School principals of the control schools were contacted via email. The email was followed by a telephone call one week later. Twenty-seven schools (three per region) were invited to participate, and of these, ten agreed initially. Unfortunately, midway through the initial data collection phase, one control school decided not to participate, leaving a total of nine control schools – two from Auckland, two from Hawke’s Bay, and one each from the remaining five regions (Waikato, Coromandel, Bay of Plenty, South Canterbury, and Southland). All Year 6 to Year 8 students from the nine control schools were invited to participate in the research project. They were initially informed about the research by the principal of the school and by their classroom teacher in early February 2013. At that time, the project was explained, and Parental Information Sheets (P-PIS), Parental NCFs, Student Information Sheets (S-PIS), and Student ASFs were distributed by teachers. Teachers agreed to administer the questionnaire to their children at both time points, and to demonstrate their willingness to participate in the administration process, they signed and returned Consent Forms (CF) and Confidentiality Agreements (CA) to the researcher. At this time, the principals also signed a CF for Principals and an Assurance Form (AF). See Appendices P-S for copies of the Teacher CF, Teacher CA, Principal CF and Principal AF for Control School participants.

### **6.3 Data Collection**

A questionnaire administration booklet was created and distributed to Kiwi Can leaders ( $n = 30$ ) and teachers ( $n = 9$ ) to ensure consistency of data collection procedures across all schools. Time 1 (i.e., pre-test or baseline) data collection began in late February 2013. In all schools, the survey was read aloud by either the Kiwi Can leader (Kiwi Can school) or the classroom teacher (control school), a process that took approximately 40 minutes. On completion, children put their questionnaire into an envelope that was then mailed by the teacher or leader in a pre-addressed postpaid bag to the researcher. Upon arrival the questionnaires were immediately assigned a unique identification number.

Data collection at Time 2 (i.e., post-test) was identical to Time 1. All Time 2 questionnaires were matched using student names and the unique identification numbers from Time 1. New participants were given new identification numbers.

#### **6.4 Data Entry and Screening**

Data entry was carried out in April and May 2013 (Time 1), and in November and December 2013 (Time 2). At the end of the data entry period a randomly generated list of 5% of the original Time 1 ( $n = 87$ ) and Time 2 ( $n = 76$ ) surveys (<http://www.randomizer.org/form.htm>) were checked for accuracy. At Time 1, there were no discrepancies identified and, at Time 2, only one error was detected. The results of this verification process enabled the researcher to confidently move to the data screening and cleaning phase of the analysis.

During the data screening process, all questionnaires were inspected manually for evidence of inaccurate and/or incomplete responding. No surveys were identified as having mono-numeric responses or obvious pictorial features. The PYD and CC data sets were then treated separately for missing values analyses (MVA). Questionnaires with more than 10% missing values (i.e., 5 in the PYD, and 3 in the CC sections) were identified and removed from the data set. At Time 1 there were 19 PYD cases and 27 CC cases that had >10% missing responses. Therefore, the 46 cases with >10% missing data were deleted from the Time 1 data set. Two tests were then conducted in order to assess whether the balance of missing data were missing completely at random. The first test, a chi-square statistic (Little's MCAR Test), assessed whether the data met this assumption (Tabachnick & Fidell, 2007). The second test, an ANOVA, using Hochberg's GT2 post hoc test, determined whether there were statistically significant differences in the mean responses for surveys with no missing data (0%), 1-10% missing data and >10% missing data. Hochberg's GT2 was chosen as the test statistic, given the large differences in sample size between the three groups (Field, 2013b). Results of the Little's MCAR Test and the ANOVAs were both non-significant at  $p > .05$  for the PYD and CC data sets. This provided support for the hypothesis that the data were missing completely at random (i.e., no identifiable pattern existed for the missing data). This rendered a data set of 1697 PYD surveys and 1689 CC surveys for Time 1.

The same process was undertaken for Time 2. At Time 2, 12 PYD surveys and 17 CC surveys were initially excluded for having >10% missing data. Little's MCAR and ANOVA tests were run and were again non-significant at  $p > .05$ . This produced data sets with 1510 PYD surveys and 1505 CC Time 2 surveys. All four data sets (i.e., Time 1 PYD, Time 1 CC, Time 2 PYD, and Time 2 CC) had remaining missing values imputed using the EM procedure. Once the EM procedure was complete, all newly imputed values were checked for accuracy. Any values  $< 1$  or  $> 6$  were rounded to the nearest whole value to retain the integrity of the original response scale. A matched data set was extracted from the Time 1 and Time 2 files, resulting in 1235 *matched* PYD cases and 1224 matched CC cases, with 1219 students providing complete and matched PYD and CC responses. This final, composite data set served as the main dataset of all subsequent analysis.

#### **6.4.1 Participants.**

The final data set had 1219 complete questionnaire sets from Year 6-8 children (Table 26). Of those, 299 attended experienced Kiwi Can schools, 464 were in new Kiwi Can schools, and 456 attended control schools. In this sample, approximately half were female ( $n = 613$ ), and just under half (47.0%) were in Year 7. The remaining children were in Years 6 and 8. Most children identified themselves as either Pakeha/European New Zealanders ( $n = 519$ ) or Māori ( $n = 455$ ). Very few self-identified as being of Pasifika, Other, or Asian ethnicity.

Table 26

*Frequency Counts for Student Participants in Kiwi Can Outcome Evaluation by School Condition, Sex, Year Level, and Ethnicity*

Variable	Time 1		Time 2		Matched T1-T2	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
School Condition						
Experienced	535	31.2	473	31.1	299	24.5
New	607	35.4	518	34.0	464	38.1
Control	574	33.4	531	34.9	456	37.4
Sex						
Girl	853	49.7	757	49.7	613	50.3
Boy	863	50.3	765	50.3	606	49.7
Year Level						
Year 6	505	29.4	462	30.4	371	30.4
Year 7	824	48.0	726	47.7	573	47.0
Year 8	387	22.6	334	21.9	273	22.4
Ethnicity						
Pakeha/NZ European	677	39.5	608	40.0	519	42.4
Māori	675	39.3	590	38.8	455	37.4
Pasifika	190	11.1	171	11.2	116	9.6
Asian	55	3.2	53	3.5	46	3.8
Other	119	6.9	99	6.5	83	6.8

## 6.5 Data Analysis

All variables were then assessed for normality. Negatively worded items in the PYD and/or CC scales were transformed. One was from the Likeability scale (L3), and the other three formed the Social-ingenuousness scale (SI1, SI2, and SI3). The item scores were reverse scored according to the following formula:

$$\text{New value} = (\text{Score} * (-1)) + (\text{Maximum score} + 1)$$

A priori PYD and CC scales were then assessed for internal reliability. The underlying factor structure of the PYD and CC scales was assessed using CFA and SEM techniques. All measurement models were also tested for invariance across time (Time 1 and Time 2) and school condition (experienced Kiwi Can, new Kiwi Can, and control schools).

The relationship of the PYD and CC constructs was investigated using regression analysis within SEM for each time period, and then changes across time were analysed using SEM. Mean scores for social development and classroom climate constructs (subscales) were computed by finding the average response for all items

contributing to the factor at Time 1 and Time 2, and mean differences compared between groups of children by school condition (experienced Kiwi Can, new Kiwi Can, and control), sex, year level, and ethnicity. Mean differences within time were explored using MANOVA tests with composite subscale means.

All MANOVA output was subjected to Box's test of equality of covariance matrices. The results for the baseline mean data for PYD and CC were significant ( $p < .05$ ), so Q-Q plots and covariance matrices were constructed. The plots were explored and matrices investigated. In each case the plots demonstrated mild kurtosis ( $< 7$ ) and larger groups within each factor had larger variances than the smaller groups. Therefore, homogeneity of variance was assumed, and the Pillai-Bartlett trace and Games-Howell test statistics were selected for use.

The effect of school condition, age, year level, and ethnicity on PYD and CC means was then explored at baseline and across time. Within time, a multiple analysis of variance technique MANOVA was used to identify significant differences and main effects. Across time, significant differences and main effects for PYD and CC means were explored using a one-way analysis of variance with repeated measures (RM-ANOVA) approach. All PYD and CC items were somewhat negatively skewed, but none were beyond the threshold of acceptability for skewness ( $> 2$ ) or for kurtosis ( $> 7$ ).

## **6.6 Results - Positive Social Development (PYD) Data**

The original PYD factor structure had four second-order (Caring/compassion, Competence, Connection and Character) and seven first-order PYD factors (Perspective Taking, Empathic Concern, Social Rules/Politeness, Connection to School, Connection to Peers, Values Diversity, and Personal Values). Each of the first-order factors, or subscales, mapped onto one of the four second-order factors (scales).

Scale #1: Caring/Compassion – Perspective Taking (PT), Empathic Concern (EC)

Scale #2: Competence – Social Rules/Politeness (SRP)

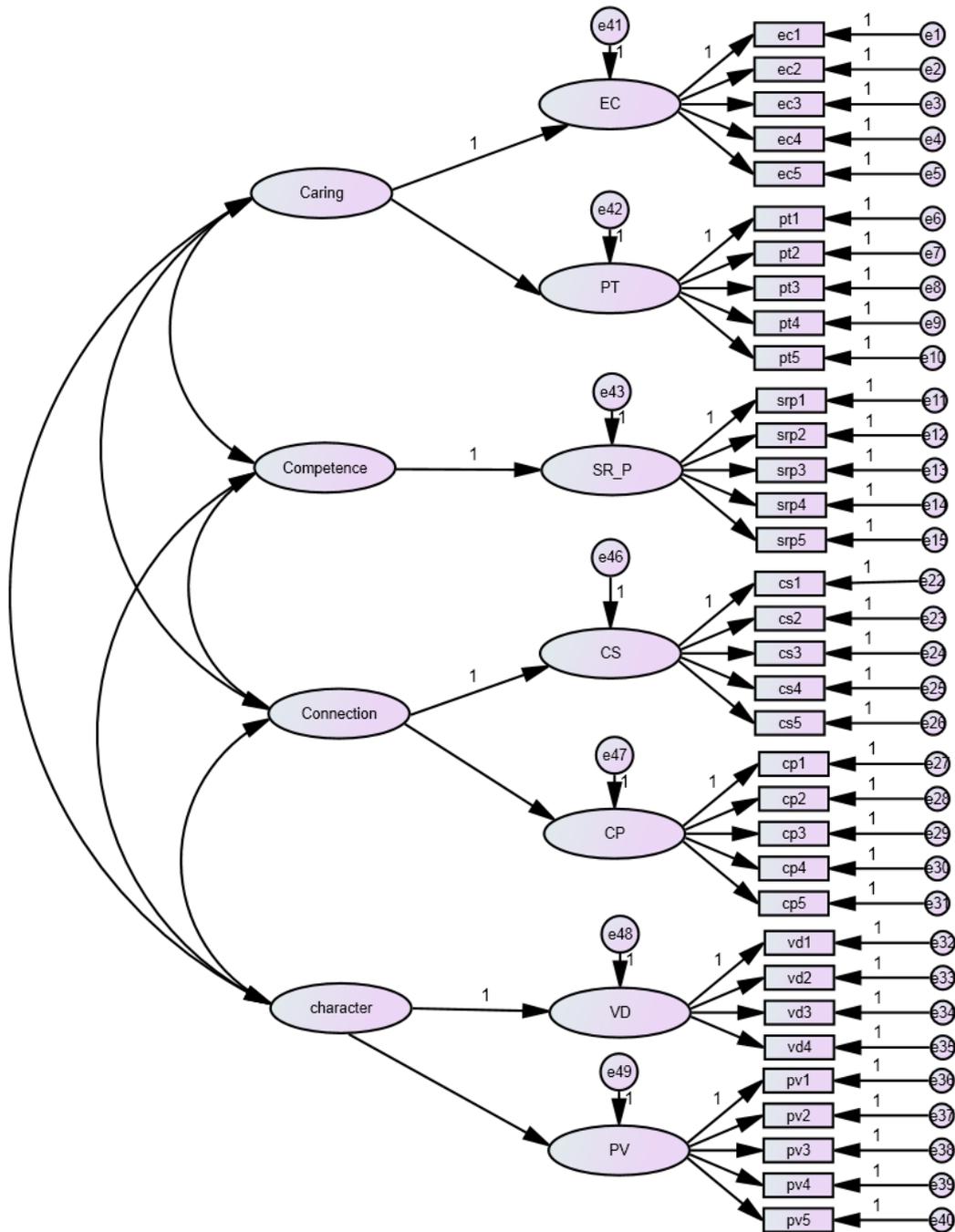
Scale #3: Connection – Connection to school (CS), Connection to peers (CP)

Scale #4: Character - Values Diversity (VD), Personal Values (PV)

Each subscale was represented by the presence of three to five items on the questionnaire.

In the testing phase, a multi-dimensional, hierarchical PYD measurement model (MM) was tested in AMOS, a statistical software programme (Figure 3). For this analysis, the maximum likelihood (ML) estimation procedure was used because the data were normally distributed, the sample size was large ( $n = 1,219$  cases), and because the response scale was long enough. All Time 1 PYD data were used for the analysis.

**Figure 3.** The PYD Model 1



*Figure 3.* A schematic of the original structural equation model with four, second order PYD scales and corresponding subscales: EC=empathic concern; PT=perspective taking; SRP =social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV=personal values.

Model 1 was inadmissible because there were two negative error variances for items associated with the Connection to School first-order factor. The model was therefore adjusted by removing the Connection to School factor and by mapping the observed variables directly onto the second-order Connection factor. Although the model solved admissibly, the goodness-of-fit statistics were poor (see Table 27), and there was very high multicollinearity between second-order factors. In fact, correlations ranged from  $r = .79$  (Connection-Character) to  $r = .99$  (Caring-Character). The average correlation was  $r = .87$ .

Table 27

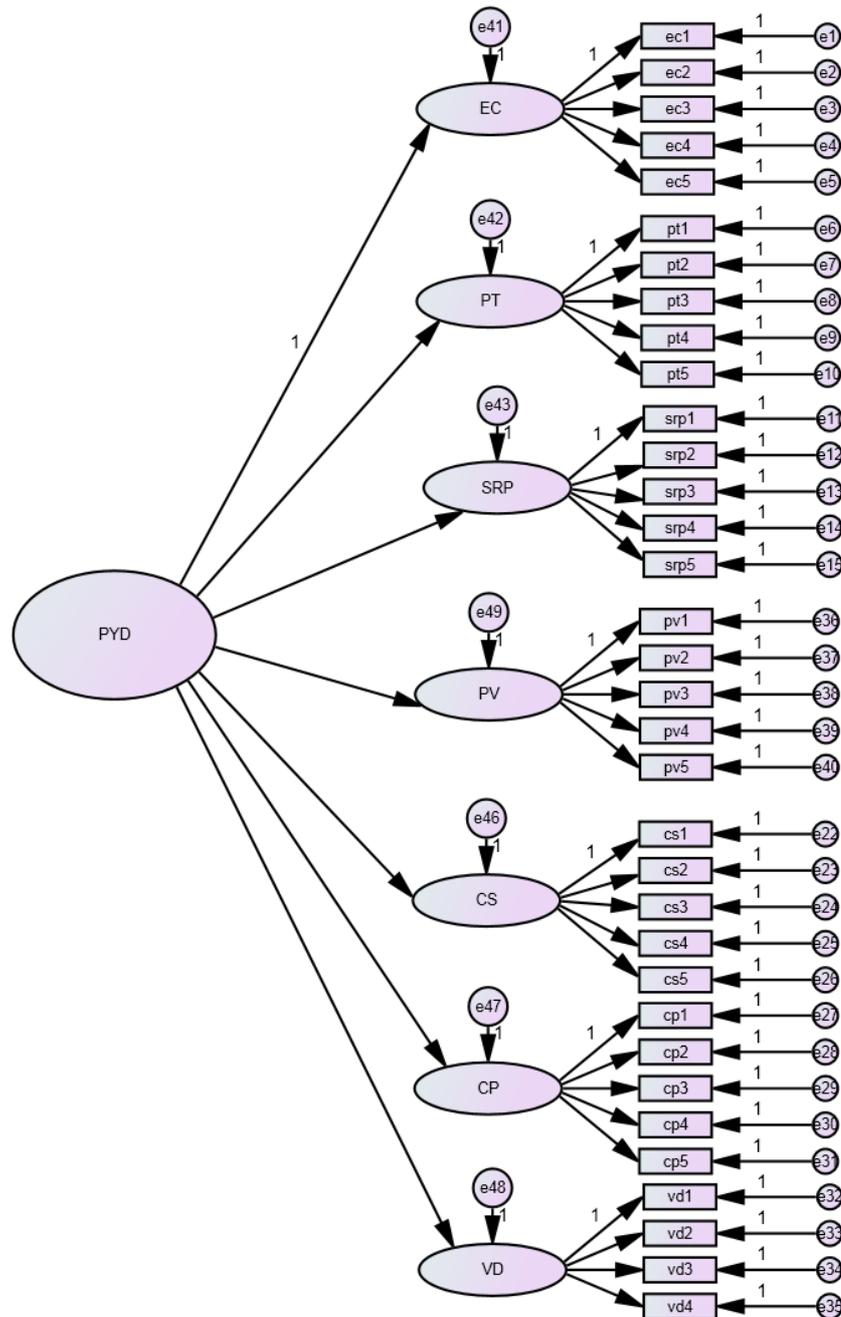
*Goodness-of-Fit Statistics for PYD Measurement Models*

Model	Characteristics	$\chi^2$	$df$	$\chi^2/df$ ( $p$ )	CFI	Gamma hat	RMSEA	SRMR
1.	Hierarchical, 4 dimensions, 8 1 <sup>st</sup> -order factors	3333.26	725	4.60 ***	.87	.90	.054	.057
1A.	As #1, minus CS 1 <sup>st</sup> -order factor	2309.82	516	4.48 ***	.87	.92	.053	.056
2.	Hierarchical, 1 dimension, 8 1 <sup>st</sup> -order factors	2618.44	520	5.04 ***	.88	.91	.058	.055
3.	7 inter-correlated 1 <sup>st</sup> -order factors	2215.98	506	4.38 ***	.90	.92	.053	.044

\*\*\* =  $p < .001$ .

This led to the design of Model 2, in which all the first-order factors were predicted by a single second-order factor (PYD) (see Figure 4). This model was admissible, but had a mixture of poor to acceptable fit statistics. Further, the standardised regression weights for three of the first-order factors were extremely high ( $\beta > .94$ ), indicating evidence of multicollinearity (see Table 27).

**Figure 4.** The PYD Model 2



*Figure 4.* Structural equation model with a single, second order PYD structural model predicted by all seven subscales: EC=empathic concern; PT=perspective taking; SRP =social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV=personal values.

Therefore, the decision was made to test an inter-correlated first-order factor model with no hierarchical second-order factor (Figure 5). Scale reliability tests (i.e., Cronbach’s alpha ( $\alpha$ )) were computed for each of the PYD scales and subscales (see Table 28). Seven of the nine PYD had adequate or sufficient reliability ( $\alpha > .70$ ) at Time 1 and Time 2 (Cronbach, 1951). Two of the PYD subscales (Likeability and Social Ingenuousness) had  $\alpha < .70$  at Time 1 or 2, suggesting issues to do with the

coherence of the items in the scale for this sample. Given that the overall integrity of the social competence factor was not threatened by the removal of these two subscales, it was decided to remove them from further analysis. This produced a PYD data set with seven subscales around four factors.

Table 28

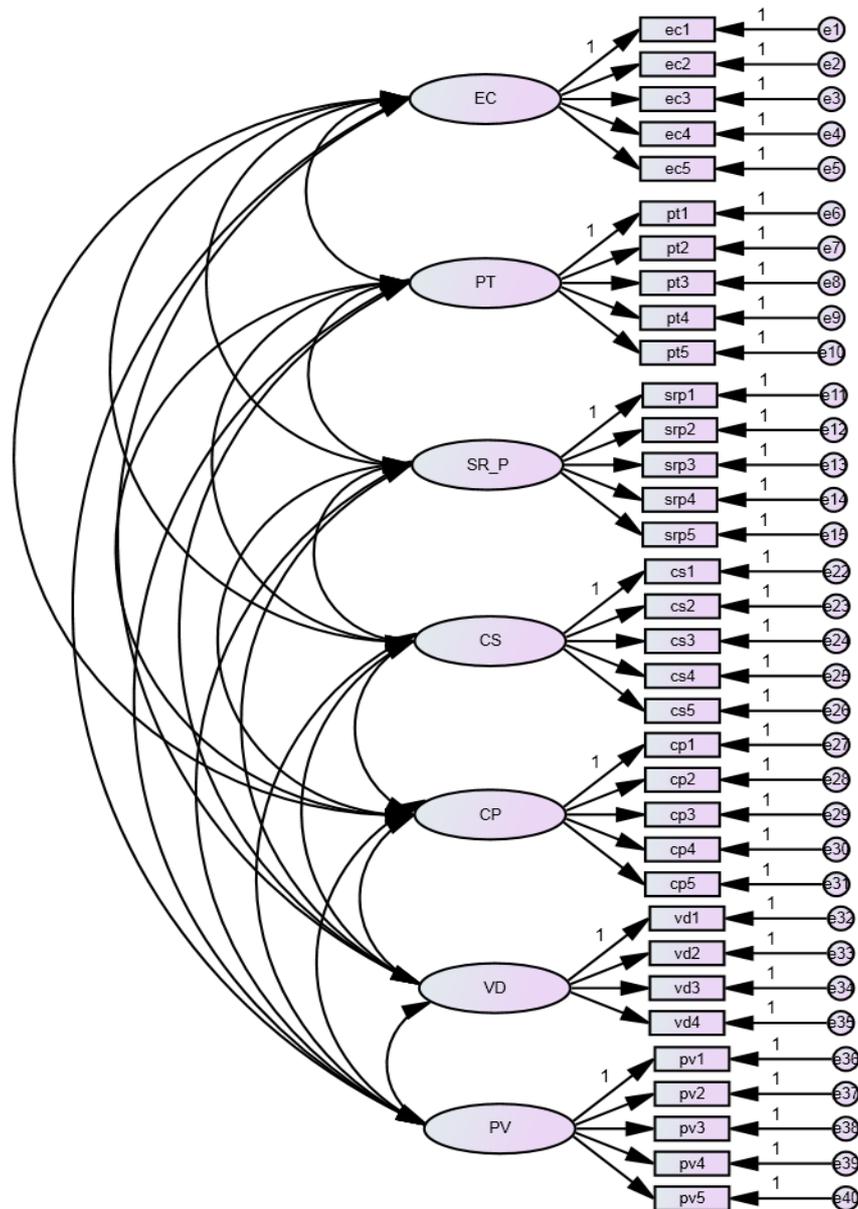
*Scale Validation for PYD Data at Time 1 and Time 2*

Subscale	Cronbach's Alpha ( $\alpha$ )	
	Time 1	Time 2
Positive Youth Development		
Empathic Concern (EC)	.81	.83
Perspective Taking (PT)	.71	.77
Social Rules/Politeness (SRP)	.78	.79
Likeability (L)	.56	.62
Social-Ingenuosness (SI)	.58	.57
Connection to School (CS)	.78	.81
Connection to Peers (CP)	.86	.88
Values Diversity (VD)	.71	.77
Personal Values (PV)	.75	.77

*Note.* Values in red fail to meet conventional standards.

The unconstrained model of seven inter-correlated factors solved and had fair to good fit statistics ( $\chi^2/df = 4.41$ ; CFI = .90; RMSEA = .053, 90% CI = [.050, .055], pclose = .028; SRMR = .044; gamma hat = .92). Further, all standardised regression weights were statistically significant at  $p < .001$  and all were  $\beta > .41$ . The correlations between latent factors were moderate to large, ranging from  $r = .42$  (Empathic Concern-Connection to Peers) to  $r = .91$  (Perspective Taking-Personal Values).

**Figure 5.** Inter-correlated PYD Model 3



*Figure 5.* Inter-correlated measurement model for seven PYD subscales: EC=empathic concern; PT=perspective taking; SRP =social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV=personal values.

### 6.6.1 Discussion of results for PYD model.

The most appropriate model for describing the relationships between PYD factors was a first-order, inter-correlated model. This did not support previous research (Jelicic et al., 2007; Lerner et al., 2005; Phelps et al., 2005) that found support for a hierarchical structure of positive youth development amongst thousands of Grade 5-12 students in the United States. In their hierarchical model, the 5C's (Connection, Competence,

Caring/compassion, Character, and Confidence) were positioned as first-order latent factors that all mapped onto a second-order PYD factor. In the present study, however, only four of the five Cs were used, and this may explain the differences. Bowers et al. (2010) recently tested the higher-order PYD model with 920 Grade 8-10 students from across 30 states in America and did not find support for the model, with inadequate fit to the data. Therefore, it can be assumed that the original, hierarchical model may not be appropriate in all settings, and that a first-order, inter-correlated measurement model (MM) is acceptable moving forward.

### 6.7 Modelling the Classroom Climate, CC, Data

The initial inter-correlated, five-factor MM was tested using the Time 1 CC data ( $n = 1,219$  cases) (see Figure 6). Model 1 had adequate fit, however there was a very high correlation between Positive Behaviour and Classroom and School Supportiveness ( $r = .98$ ).

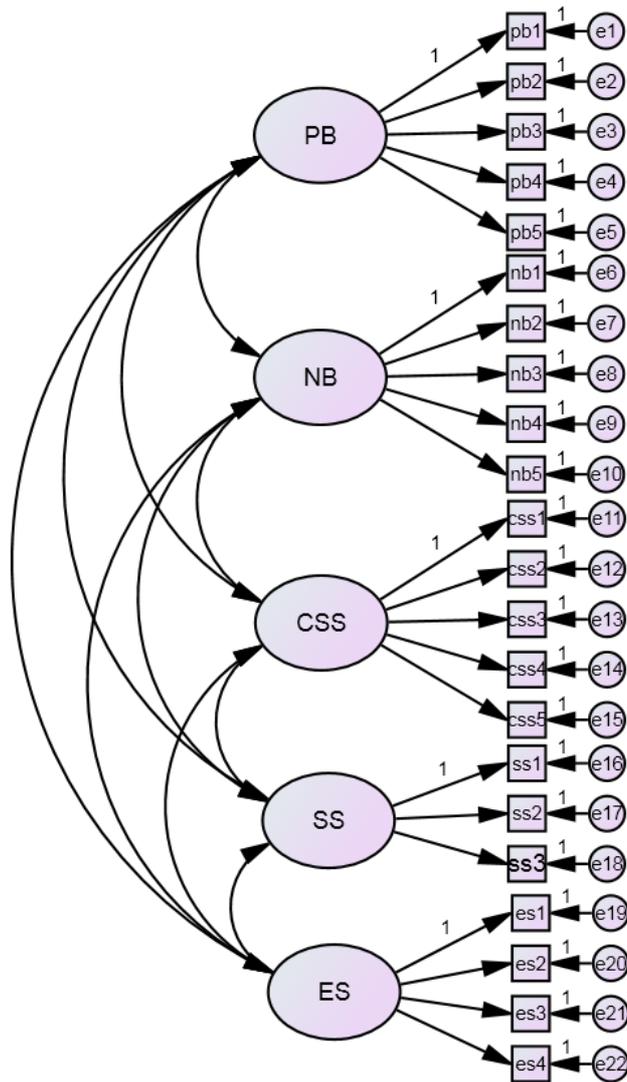
Table 29

*Goodness-of-Fit Statistics for CC Measurement Models*

Model	Characteristics	$\chi^2$	$df$	$\chi^2/df$ ( $p$ )	CFI	Gamma hat	RMSEA	SRMR
1.	Inter-correlated, first order, five factor model	825.70	199	4.15 ***	.94	.96	.051	.037
2.	Inter-correlated, first order, four factor model	860.63	203	4.24 ***	.93	.95	.051	.039

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ .

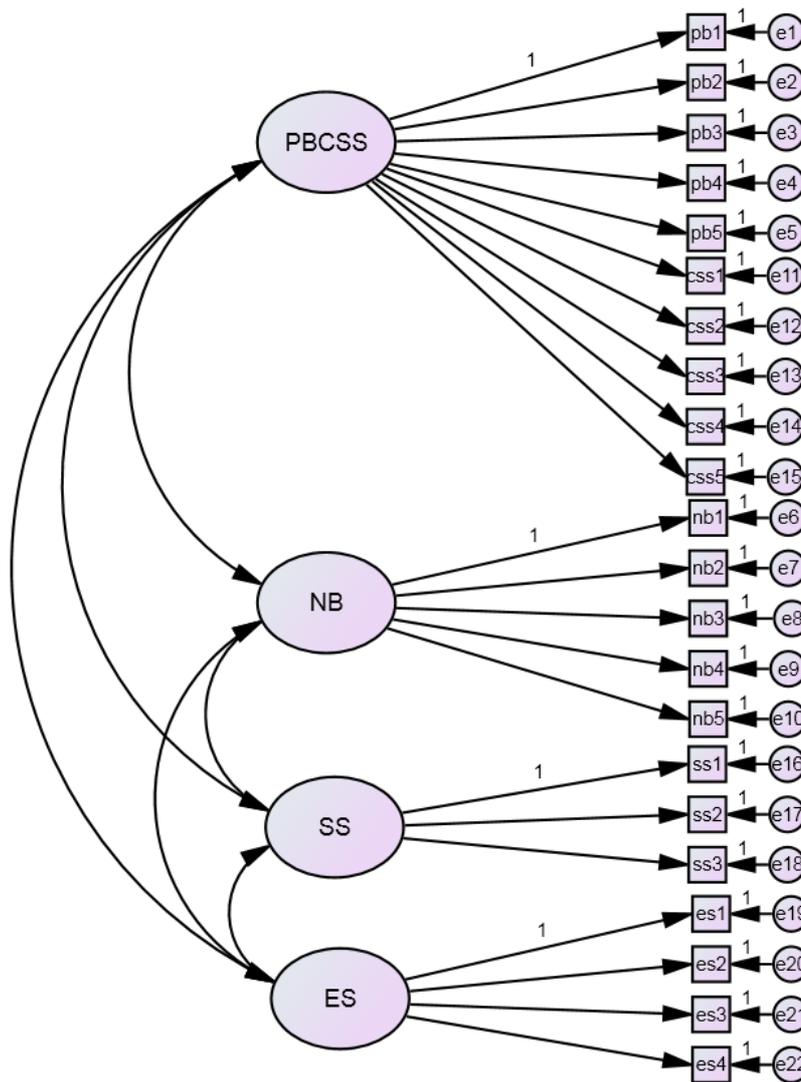
**Figure 6.** The Original Five-Factor Measurement Model for CC Data



*Figure 6.* A copy of the original five-factor classroom climate measurement model. PB=positive behavior; NB=negative behavior; CSS=classroom and school supportiveness; SS=safety at school; ES=enjoyment of school.

Therefore, these two factors were merged into a single factor, Positive Behavior and Classroom (and School) Supportiveness (PBCSS), and a four-factor, inter-correlated model was tested (see Figure 7). Model 2 was admissible and had acceptable fit (see Table 29). Further, the factor inter-correlations ranged from weak to strong, supporting independence of the four factors (correlations ranged from  $r = .29$  to  $.82$ ).

**Figure 7.** Four-factor, Inter-correlated CC Model



*Figure 7.* A schematic of the four-factor, inter-correlated classroom climate measurement model. PB=positive behavior; NB=negative behavior; CSS=classroom supportiveness; SS=safety at school; ES=enjoyment of school.

Given the acceptable fit and invariance across time in this model, it was retained as the final MM for the CC construct. Scale reliability tests (i.e., Cronbach's alpha ( $\alpha$ )) were computed for each of the CC scales and subscales (see Table 30). All of the CC scales had adequate or sufficient reliability ( $\alpha > .70$ ) at Time 1 and Time 2.

Table 30

*Scale Validation for PYD Data at Time 1 and Time 2*

Subscale	Cronbach's Alpha ( $\alpha$ )	
	Time 1	Time 2
Classroom Climate		
Positive Behavior (PB)	.76	.80
Negative Behavior (NB)	.75	.78
Classroom and School Supportiveness (CSS)	.82	.85
Safety at School (SS)	.72	.78
Enjoyment of School (ES)	.78	.80

**6.7.1 Discussion of results for CC model.**

The final MM for the climate of the classroom is very similar to the original model proposed and tested by Ding et al. (2011) with Canadian youth. Interestingly, the model generated with Year 6-8 children in New Zealand does have one unique feature in that the Positive Behaviour and Classroom and School Supportiveness scales have been amalgamated. From a theoretical perspective, the joint scale is appropriate given that the two scales have items that relate directly to positive, prosocial attitudes and behaviours within the school classroom (i.e., "Students in my class help each other learn"). More specifically, the two scales include statements about the degree to which students demonstrate caring, competence, and high moral character. The Positive Behaviour scale focuses on issues of equity, fairness, and adherence to social rules. The Classroom and School Supportiveness scale focuses more on items related to demonstrable care and compassion. Within the New Zealand sample, children responded to both scales in a highly consistent manner suggesting that they did not distinguish between a commitment to fairness/equity and caring/compassion. Although this is an interesting finding, the research methods used in this study do not allow evaluation of possible causes for the result.

**6.8 The Relationship Between PYD and CC**

Multigroup invariance testing was used to establish that responses to both models were equivalent between the two times of administration and across the three experimental conditions. Both the PYD and CC MMs demonstrated strict invariance for both time and school condition as the sequential change in CFI was within the acceptable range ( $\Delta\text{CFI} < .01$ ) (see Table 31).

Table 31

*Results of Invariance Testing for Final PYD and CC Measurement Model*

Model & Invariance Test	Time		School Condition	
	CFI	$\Delta$ CFI	CFI	$\Delta$ CFI
Final PYD MM				
Unconstrained	.900	—	.901	—
Measurement equivalence	.899	.001	.900	.001
Structural covariances	.898	.001	.898	.002
Measurement residuals	.894	.004	.896	.002
Latent Means Structures	.897	.003	.898	.002
Final CC MM				
Unconstrained	.929	—	.930	—
Measurement invariance	.927	.002	.928	.002
Structural covariances	.919	.008	.924	.004
Measurement residuals	.914	.005	.923	.009
Latent Means Structures	.923	.009	.927	.004

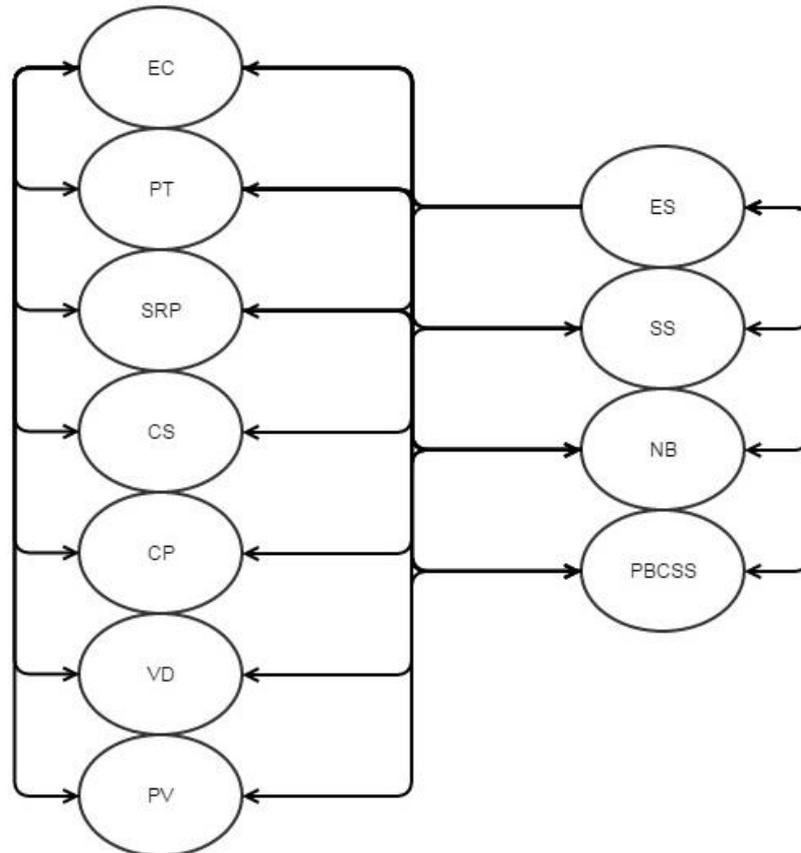
**6.9 Designing and Testing the SEM**

Multiple structural models were then created to test alternative plausible theories of how PYD and CC might be related. Model 1 was a simple inter-correlated model in which all factors of CC were correlated with PYD, on the assumption that these two constructs mutually influence each other without any causal relationship. Model 2 was a causal model in which the PYD factors were predictors of the CC factors, on the assumption that, according to PYD theory, the accumulation of assets and general process of positive development supports prosocial behaviour and thus contributes to a more prosocial classroom climate. Model 3 was a reverse direction predictive model in which the CC factors were predictors of the PYD factors. The logic for this model is that positive classroom environments would act as an external asset that individual children accrued, and that over time would lead to the emergence of PYD outcomes. It must be noted that positive classroom climate is a necessary, but on its own insufficient condition for producing PYD, though it may play a role in the emergence of PYD factors. Model 4 was a hierarchical model for both PYD and CC based on previous studies with similarly aged children in the United States (Jelicic et al., 2007; Lerner et al. 2005; Phelps et al., 2009). Model 4A positioned PYD as the predictor of CC, while Model 4B reversed the direction.

### 6.9.1 Model 1: Inter-correlated PYD and CC.

The inter-correlated model was created by merging the MMs for PYD and CC into a single SEM. In this model, all seven, Time 1 PYD and four CC latent variables were correlated with one other (see Figure 8 for a schematic of the inter-correlated Model 1).

**Figure 8.** Model 1: An Inter-correlated PYD to CC Model



*Figure 8.* Schematic of Model 1: Inter-correlated PYD-CC. EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV=personal values; ES= enjoyment of school, SS= safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness.

This model was admissible and had acceptable fit statistics (see Table 32).

Table 32

*Goodness-of-Fit Statistics for PYD and CC Structural Equation Models*

Model	Characteristics	$\chi^2$	<i>df</i>	$\chi^2/df$ ( <i>p</i> )	CFI	Gamma hat	RMSEA	SRMR
1.	Inter-correlated model with 7 PYD and 4 CC factors	7,632.60	1,429	5.34 ***	.90	.92	.042	.038
2.	Regression model PYD predicts CC with all factor pathways	4,639.92	1,429	3.25 ***	.89	.97	.043	.041
2A.	Regression model PYD predicts CC with 14 deleted pathways	1,943.47	668	2.91 ***	.93	.98	.040	.036
3.	Regression model CC predicts PYD with all factor pathways	7,632.60	1,429	5.34 ***	.90	.92	.042	.045
3A.	Regression model CC predicts PYD with insignificant pathways deleted	7,113.32	1,333	5.34 ***	.90	.92	.042	.038
4.	Hierarchical regression model PYD predicts CC	5,495.23	1,472	3.73 ***	.86	.95	.047	.055
4A.	Hierarchical regression model CC predicts PYD	5,495.23	1,472	3.73 ***	.86	.95	.047	.055
5.	Uncorrelated structural Time1 and Time 2 PYD—CC Models	19,533.15	6,080	3.21 ***	.80	.53	.043	.134
6.	Autoregressive inter-correlated PYD -- CC model	17,298.00	6,069	2.85 ***	.83	.88	.039	.053
7.	Autoregressive inter-correlated PYD – CC model with insignificant pathways removed	17,296.28	6,067	2.85 ***	.83	.93	.039	.053
8.	Autoregressive cross-lagged model PYD1 to CC2	16,962.86	6,028	2.81 ***	.84	.94	.039	.051

Model	Characteristics	$\chi^2$	<i>df</i>	$\chi^2/df$ ( <i>p</i> )	CFI	Gamma hat	RMSEA	SRMR
8-2.	Autoregressive cross-lagged model PYD1 to CC2 trimmed	16,983.96	6,046	2.81 ***	.84	.94	.039	.051
8A.	Autoregressive cross-lagged model CC1 to PYD2	17,072.76	6,047	2.82 ***	.84	.94	.039	.051
8A-2.	Autoregressive cross-lagged model CC1 to PYD2 trimmed	17,092.42	6,056	2.82 ***	.84	.94	.039	.051

\*\*\* =  $p < .001$ .

All variables except for Negative Behaviour co-varied positively (see Table 33 for a list of covariances as well as baseline means and standard deviations for the PYD and CC subscales at Time 1). This was not surprising given that Negative Behaviour measured incidence of negative antisocial behavior, while the remaining variables measured incidence of positive prosocial attitudes and behaviours.

Table 33

*Covariances for Inter-correlated PYD and CC Structural Equation Models*

Factors	PYD							CC			
	EC	PT	SRP	CS	CP	VD	PV	ES	SS	NB	PBCSS
Mean (SD)	5.09 (.86)	4.59 (.88)	5.31(.68)	5.05(.79)	4.82 (.95)	4.84(.84)	5.00(.80)	4.97(1.07)	4.94(.90)	3.35(1.11)	4.73(.80)
EC	—										
PT	.25***	—									
SRP	.18***	.19***	—								
CS	.22***	.23***	.21***	—							
CP	.21***	.24***	.21***	.44***	—						
VD	.22***	.22***	.15***	.22***	.22***	—					
PV	.22***	.25***	.19***	.27***	.25***	.21***	—				
ES	.26***	.27***	.23***	.51***	.50***	.27***	.31***	—			
SS	.20***	.22***	.20***	.35***	.40***	.19***	.22***	.45***	—		
NB	-.07***	-.08***	-.08***	-.12***	-.18***	-.03***	-.08**	-.15***	-.16***	—	
PBCSS	.26***	.29***	.23***	.43***	.54***	.24***	.28***	.59***	.52***	-.25 ***	—

*Note.* EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV=personal values; ES=enjoyment of school, SS=safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness.

\*\*\* =  $p < .001$ ; \*\* =  $p < .01$ .

Lastly, analyses of invariance using the multiple group analysis function indicated that it demonstrated configural, metric, and scalar invariance across time and school condition ( $\Delta\text{CFI} < .01$ ) (see Table 34).

Table 34

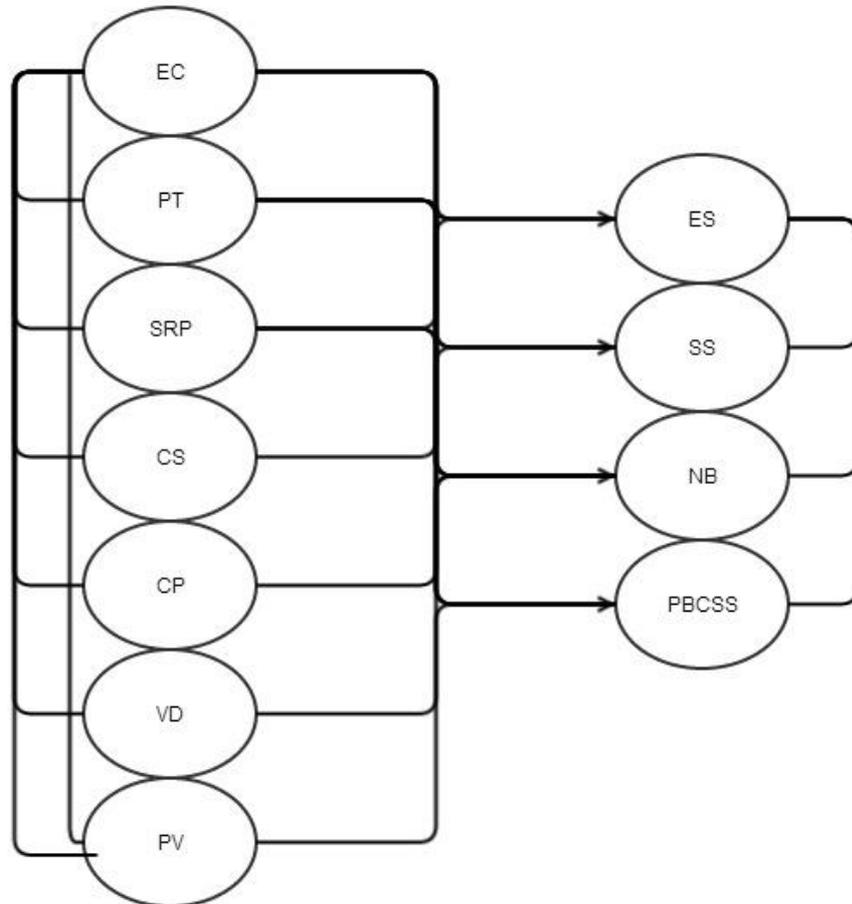
*Results of Invariance Testing for Inter-correlated PYD-CC Structural Equation Model*

Invariance Test	<u>Time</u>		<u>School Condition</u>	
	CFI	$\Delta\text{CFI}$	CFI	$\Delta\text{CFI}$
Configural invariance	CFI=.887	—	CFI=.888	—
Measurement weights	CFI=.885	.002	CFI=.887	.001
Measurement intercepts	CFI=.882	.003	CFI=.884	.003
Structural covariances	CFI=.881	.001	CFI=.882	.002
Measurement residuals	CFI=.878	.003	CFI=.880	.002

**6.9.2 Model 2: PYD to CC**

Model 2 had regression pathways from each of the seven PYD factors (Empathic Concern, Perspective Taking, Social Rules/Politeness, Connection to School, Connection to Peers, Values Diversity, Personal Values) to the four CC factors (Positive Behaviour and Classroom Supportiveness, Negative Behaviour, Safety at School, and Enjoyment of School). The original PYD-CC SEM model had regression pathways from each of the seven latent PYD factors to each of the four, latent CC factors (Figure 9). The unconstrained model was admissible and had good model fit (see Table 32).

**Figure 9.** Model 2: PYD to CC Structural Equation Model



*Figure 9.* Schematic PYD to CC structural model. EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV= personal values; ES=enjoyment of school, SS=safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness.

A number of the regression weights were statistically non-significant, but after deleting 14 pathways, Model 2A had improved fit statistics. See Table 32 for goodness-of-fit statistics and Table 35 for regression weight values and associated significance for the model.

Table 35

*Regression Weights for Model 2A: PYD predicts CC*

CC Factors		PYD Factors	Standardized Regression Weight $\beta$	<i>p</i>
PBCSS	<---	EC	.15	**
PBCSS	<---	PT	.34	***
SS	<---	SRP	.77	***
PBCSS	<---	CS	.20	***
ES	<---	CS	.97	***
PBCSS	<---	CP	.52	***
SS	<---	CP	.37	***

*Note.* EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV= personal values; ES=enjoyment of school, SS=safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness.

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ .

When the characteristics of the model were explored in detail, it became clear that both of the Character subscales (Personal Values and Values Diversity) were not at all predictive of scores for classroom climate variables. Alternatively, individual scores in caring and compassion (Empathic Concern, Perspective Taking) positively predicted variability in the positive behavior (Positive Behaviour and Classroom Supportiveness) and safety subscales (Safety at School). Scores for Safety at School were also positively predicted by scores for social competence (Social Rules/Politeness). This indicated that individuals who were more competent and caring were likely to report experiencing a safe and positive school environment. Not surprisingly, individual scores for connection to the school environment (Connection to School) were predictive of positive behavior at school (Positive Behaviour and Classroom Supportiveness), while scores for connection to one's peers (Connection to Peers) predicted variability in the safety and positive behaviour dimensions. None of the individual level variables predicted scores for negative or antisocial behavior, nor did they explain much of the variance in scores for the enjoyment of school factor (Enjoyment of School).

Model 2A had configural, measurement, and scalar invariance for both Time 1 and Time 2, and for school condition (see Table 36).

Table 36

*Results of Invariance Testing for the Final PYD Predicts CC Structural Equation Model*

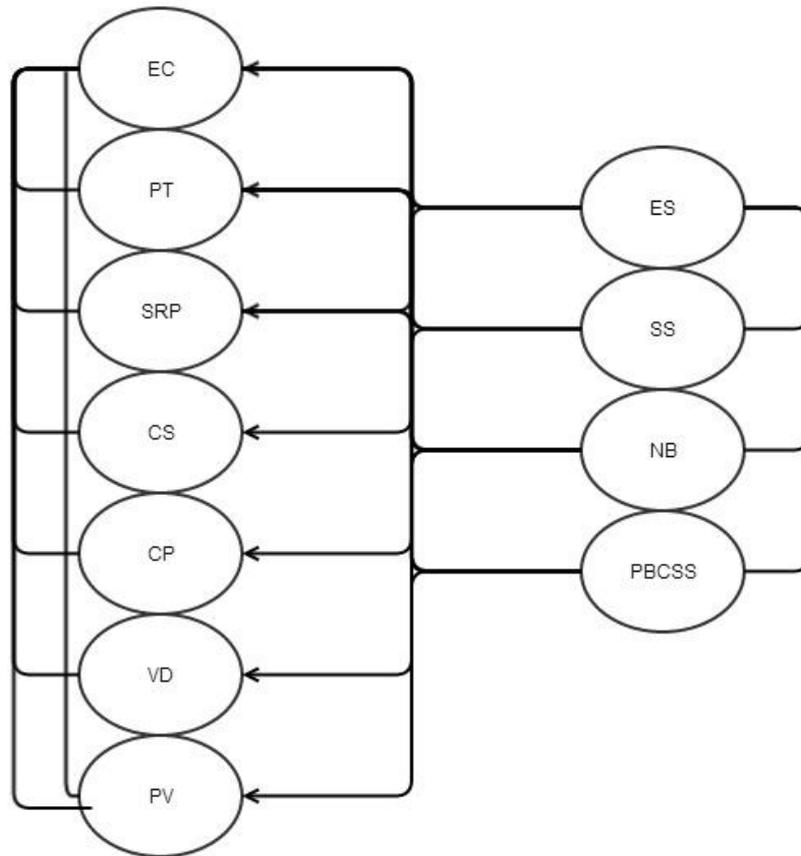
Invariance Test	Time		School condition	
	CFI	ΔCFI	CFI	ΔCFI
Configural invariance	CFI=.888	—	CFI=.888	—
Measurement weights	CFI=.887	.001	CFI=.887	.001
Measurement intercepts	CFI=.884	.003	CFI=.885	.002
Structural weights	CFI=.884	.000	CFI=.885	.000
Structural residuals	CFI=.883	.001	CFI=.884	.001
Measurement residuals	CFI=.880	.003	CFI=.881	.003

Given that the model demonstrated invariance, had adequate fit, and was theoretically sound, Model 2A was retained as the best representation of PYD factors as predictors of CC.

**6.9.3 Model 3: CC to PYD**

Model 3, in which classroom climate predicted positive social development in youth, was tested and had adequate fit statistics. See Figure 10 for a full schematic diagram of Model 3.

**Figure 10.** Model 3: CC to PYD Model



*Figure 10.* Model 3 Schematic CC predicting PYD. EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV= personal values; ES= enjoyment of school, SS=safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness.

Again, there were a number of statistically non-significant pathways. In fact, Negative Behaviour was not predictive of any PYD factors, and Safety at School was predictive of just two PYD factors (Social Rules/Politeness and Connection to School). After removing those pathways, Model 3A demonstrated slightly better fit statistics (see Table 33). According to Model 3A, individual scores for positive behavior within the classroom (Positive Behaviour and Classroom Supportiveness) and general enjoyment in the school environment (Enjoyment of School) significantly predicted the variability in scores for Empathic Concern (Caring), Perspective Taking (Caring), Social Rules/Politeness (Competence), Connection to School (C0nnection), Connection to Peers (Connection), Personal Values, and Values Diversity (Character). Of the two scales, Positive Behaviour and Classroom and School Supportiveness predicted a higher degree of variability across all seven subscales – however, it was

moderate at best. In fact, Positive Behaviour and Classroom Supportiveness and Enjoyment of School scores predicted between 1.6% and 36.7% of the variability in PYD subscale scores. When scores for safety (SS) were explored, they too were predictive of individual PYD scores, however to a lesser degree. Safety at School scores were significantly predictive of only two subscales – Social Rules/Politeness and Connection to School. The self-reported scores for negative, or antisocial, attitudes and/or behaviours were not significantly predictive of scores for any of the positive social development-related outcomes (see Table 37 for all significant regression weights).

Table 37

*Significant Regression Weights for Model 3A: CC Predicts PYD*

PYD Factors	CC Factors	Standardized Regression Weight ( $\beta$ )	<i>p</i>
EC	← PBCSS	.26	***
EC	← ES	.12	***
SRP	← PBCSS	.28	***
SRP	← SS	.22	***
SRP	← ES	.14	***
PT	← ES	.07	**
CS	← PBCSS	.17	***
CS	← SS	.15	***
CS	← ES	.38	***
CP	← PBCSS	.64	***
CP	← ES	.13	***
VD	← PBCSS	.20	***
VD	← ES	.17	***
PV	← PBCSS	.22	***
PV	← ES	.21	***

*Note.* EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV= personal values; ES= enjoyment of school, SS=safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness; \*\*\*= $p < .001$ ; \*\* =  $p < .01$ .

Model 3A demonstrated invariance across school condition and time. See Table 38 for measures of CFI at the configural, measurement, and scalar levels.

Table 38

*Results of Invariance Testing for the Final CC Predicts PYD Structural Equation Model*

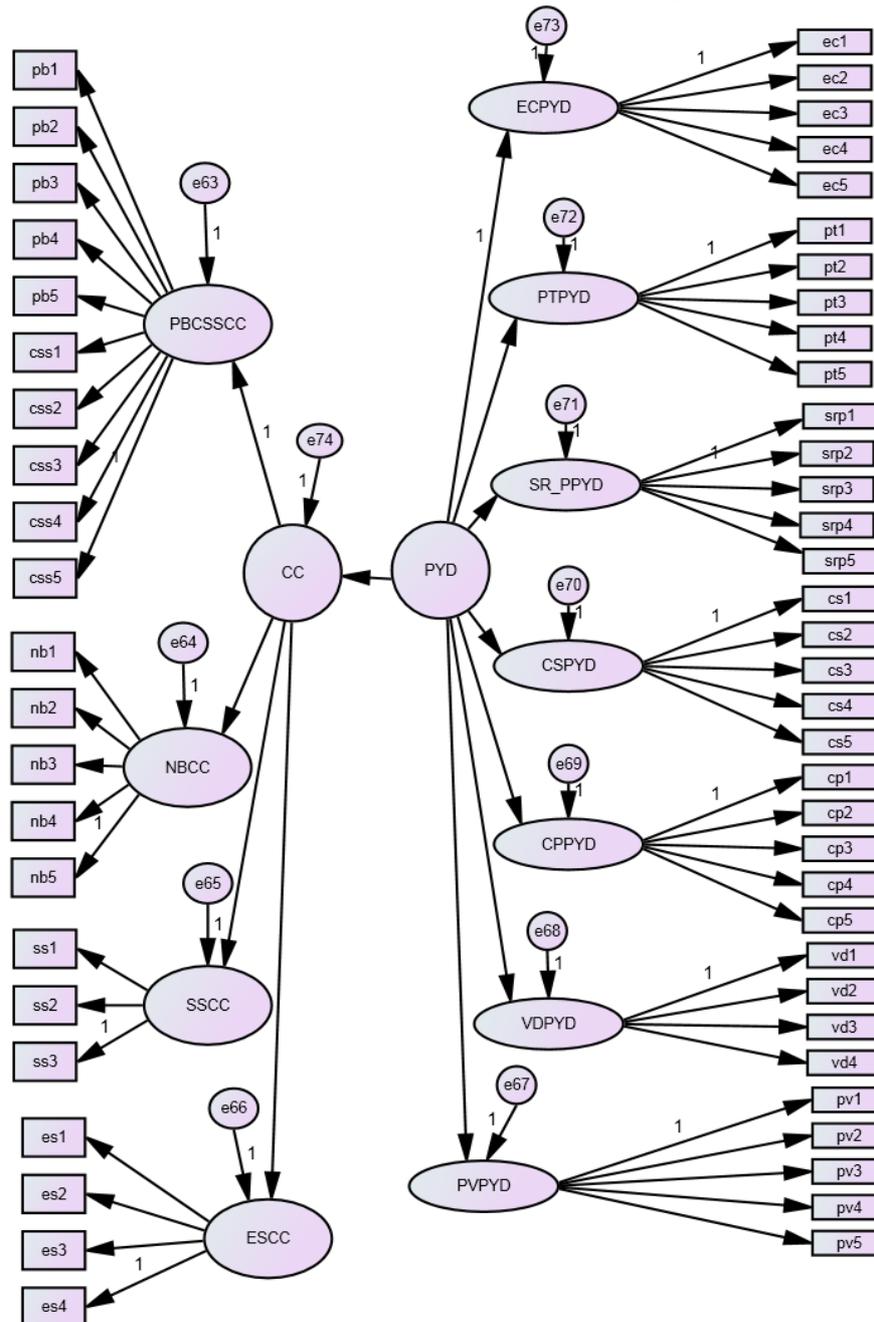
Invariance Test	Time		School condition	
	CFI	$\Delta$ CFI	CFI	$\Delta$ CFI
Configural invariance	CFI = .891	—	CFI = .893	—
Measurement weights	CFI = .890	.001	CFI = .892	.001
Measurement intercepts	CFI = .890	.000	CFI = .889	.003
Structural weights	CFI = .890	.000	CFI = .888	.001
Structural residuals	CFI = .889	.001	CFI = .887	.002
Measurement residuals	CFI = .886	.003	CFI = .885	.002

**Hierarchical Regression Models**

**6.9.4 Model 4: Hierarchical PYD to hierarchical CC.**

The PYD predicts CC hierarchical model had good fit statistics (Table 32). The standardised regression weight for the PYD predicts CC was  $\beta = .86$ , accounting for 73.6% of the variance in CC (see Figure 11).

**Figure 11.** Model 4: A Hierarchical PYD to CC Structural Equation Model

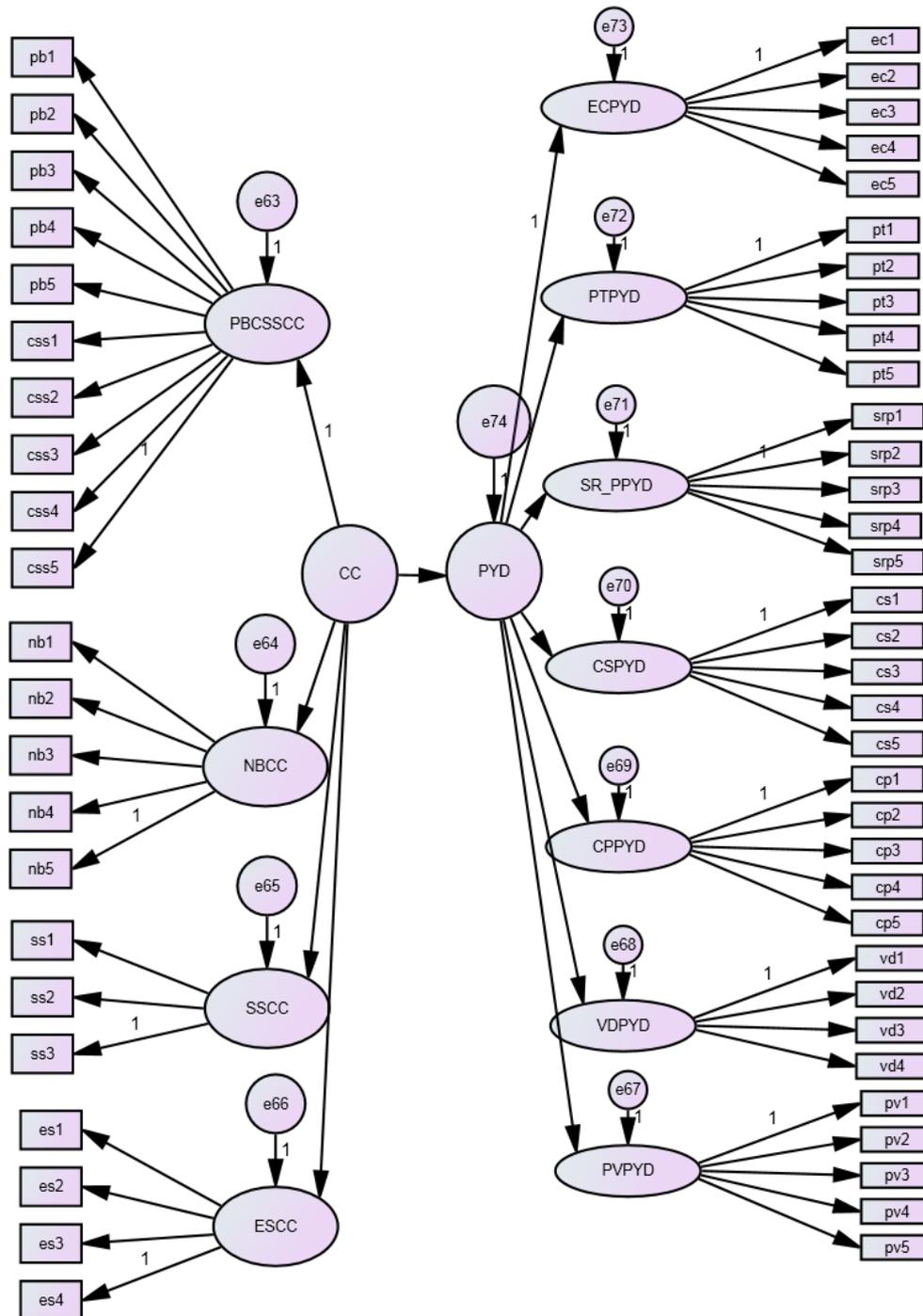


*Figure 11.* Hierarchical PYD to CC SEM. ECPYD=empathic concern; PTPYD=perspective taking; SRPPYD= social rules/politeness; CSPYD=connection to school; CPPYD=connection to peers; VDPYD=values diversity; PVPYD= personal values; ESCC= enjoyment of school, SSCC=safety at school; NBCC= negative behaviour; PBCSSCC=positive behaviour and classroom supportiveness; CC=classroom climate; PYD=positive social development in youth.

### 6.9.5 Model 4A: Hierarchical CC to hierarchical PYD.

The hierarchical or second order CC predicts PYD SEM had identical fit to Model 4A (see Figure 12). The standardised regression weight from CC predicts PYD was  $\beta = .82$ , accounting for 67.1% of the variance in PYD.

**Figure 12.** Model 4A: A Hierarchical CC to PYD Structural Equation Model



*Figure 12.* A copy of the final, second order CC → PYD SEM. ECPYD=empathic concern; PTPYD=perspective taking; SRPPYD= social rules/politeness; CSPYD=connection to school; CPPYD=connection to peers; VDPYD=values diversity; PVPYD= personal values; ESCC= enjoyment of school, SSCC=safety at school; NBCC= negative behaviour; PBCSSCC=positive behaviour and classroom supportiveness; CC=classroom climate; PYD=positive social development in youth.

Both Models 4 and 4A were strictly invariant across time and group. The CFI and  $\Delta$ CFI values for Model 4 and 4A were identical (see Table 39).

Table 39

*Results of Invariance Testing for Model 4 and Model 4A*

Invariance Test	<u>Time</u>		<u>School condition</u>	
	CFI	$\Delta$ CFI	CFI	$\Delta$ CFI
Configural invariance	CFI= .857	—	CFI= .857	—
Measurement weights	CFI= .853	.004	CFI= .853	.004
Measurement intercepts	CFI= .853	.000	CFI= .853	.000
Structural weights	CFI= .853	.000	CFI= .853	.000
Structural residuals	CFI= .852	.001	CFI= .852	.001
Measurement residuals	CFI= .849	.003	CFI= .849	.003

In order to select the preferred model between 4A and 4B, a comparison of the standardised regression weights for all latent factors to items and latent factor to latent factor was made. This was necessary because the AIC and  $\chi^2$  values for both models were identical.

Table 40 shows the standardised regression weights for the two models. The  $\beta$  values for the second order factor were stronger for Model 4A than Model 4B, and provided greater variance explained in the dependent latent factor. Hence, Model 4A was selected as the best representation of how PYD and CC scales should be related to each other.

Table 40

*Regression Weights for Hierarchical Structural Equation Model 4 and Model*

Model	Unstandardised Regression Weight ( $\beta$ )	SE	p
Model 4 PYD $\rightarrow$ CC			
CC $\leftarrow$ PYD	1.52	.11	***
PV $\leftarrow$ PYD	1.10	.09	***
VD $\leftarrow$ PYD	.99	.07	***
CP $\leftarrow$ PYD	1.36	1.00	***
CS $\leftarrow$ PYD	1.19	.08	***
SRP $\leftarrow$ PYD	1.54	.10	***
PT $\leftarrow$ PYD	1.04	1.00	***
EC $\leftarrow$ PYD	1.00		
PBCSS $\leftarrow$ CC	1.00		
NB $\leftarrow$ CC	-.35	.04	***
SS $\leftarrow$ CC	.80	.05	***
ES $\leftarrow$ CC	.89	.05	***
Model 4A CC $\rightarrow$ PYD			
PYD $\leftarrow$ CC	.44	.03	***
PV $\leftarrow$ PYD	1.10	.09	***
VD $\leftarrow$ PYD	.90	.07	***
CP $\leftarrow$ PYD	1.36	1.00	***
CS $\leftarrow$ PYD	1.19	.08	***
SRP $\leftarrow$ PYD	1.54	.10	***
PT $\leftarrow$ PYD	1.04	1.00	***
EC $\leftarrow$ PYD	1.00		
PBCSS $\leftarrow$ CC	1.00		
NB $\leftarrow$ CC	-.35	.04	***
SS $\leftarrow$ CC	.80	.05	***
ES $\leftarrow$ CC	.89	.05	***

*Note.* EC=empathic concern; PT=perspective taking; SRP=social rules/politeness; CS=connection to school; CP=connection to peers; VD=values diversity; PV= personal values; ES= enjoyment of school, SS=safety at school; NB=negative behaviour; PBCSS=positive behaviour and classroom supportiveness.

\*\*\*=  $p < .001$ .

## 6.10 Modelling Change Across Time

Having established a cross-sectional SEM that was invariant across times and groups, the next analysis investigated the relationship of Time 1 to Time 2. Four different models were tested. Model 5 posited a zero relationship between Time 1 data and Time 2 data. Not surprisingly, model fit for the uncorrelated models was poor (Table 32). Model 6 added autoregressive (AR) pathways to the model. Autoregressive paths assume that the best predictor of a response at Time 2 is the matching response at Time 1. Hence, the pathways ran from all first and second-order Time 1 PYD factors to all first and second-order Time 2 PYD factors, and from all Time 1 CC factors to all Time 2 CC factors. The fully AR model had mostly good fit indices (see Table 32). Model 7 removed the statistically non-significant pathways from Model 6,

and generated an improvement in fit. Model 7 was tested for invariance across school condition. It demonstrated configural, metric, and scalar invariance (see Table 41).

Table 41

*Results of Invariance Testing for Model 7*

Invariance Test	School Condition	
	CFI	$\Delta$ CFI
Configural invariance	CFI= .794	—
Measurement weights	CFI= .792	.002
Measurement intercepts	CFI= .789	.003
Structural weights	CFI= .789	.000
Structural covariances	CFI= .789	.000
Structural residuals	CFI= .788	.001
Measurement residuals	CFI= .785	.003

According to Model 7, all PYD and CC factors at Time 1 were predictive of scores across the same dimensions at Time 2. See Table 42 for a copy of the standardised regression weights,  $R^2$ , and effect size differences ( $f^2$ ).

Table 42

*Standardised Regression Weights for Time 1 and Time 2 PYD and CC factors*

Time 1	Time 2	Standardised regression weight ( $\beta$ )	$R^2$	Effect size differences (Cohen's $f^2$ )
PBCSS	PBCSS2	.54	.29	.41
NBCC	NBCC2	.56	.31	.45
SSCC	SSCC2	.51	.26	.35
ESCC	ESCC2	.58	.33	.49
ECPYD	ECPYD2	.63	.40	.67
PTPYD	PTPYD2	.66	.44	.79
SRPPYD	SRPPYD2	.60	.36	.56
CSPYD	CSPYD2	.61	.37	.59
CPPYD	CPPYD2	.56	.31	.45
VDPYD	VDPYD2	.61	.37	.59
PVPYD	PVPYD2	.62	.38	.61

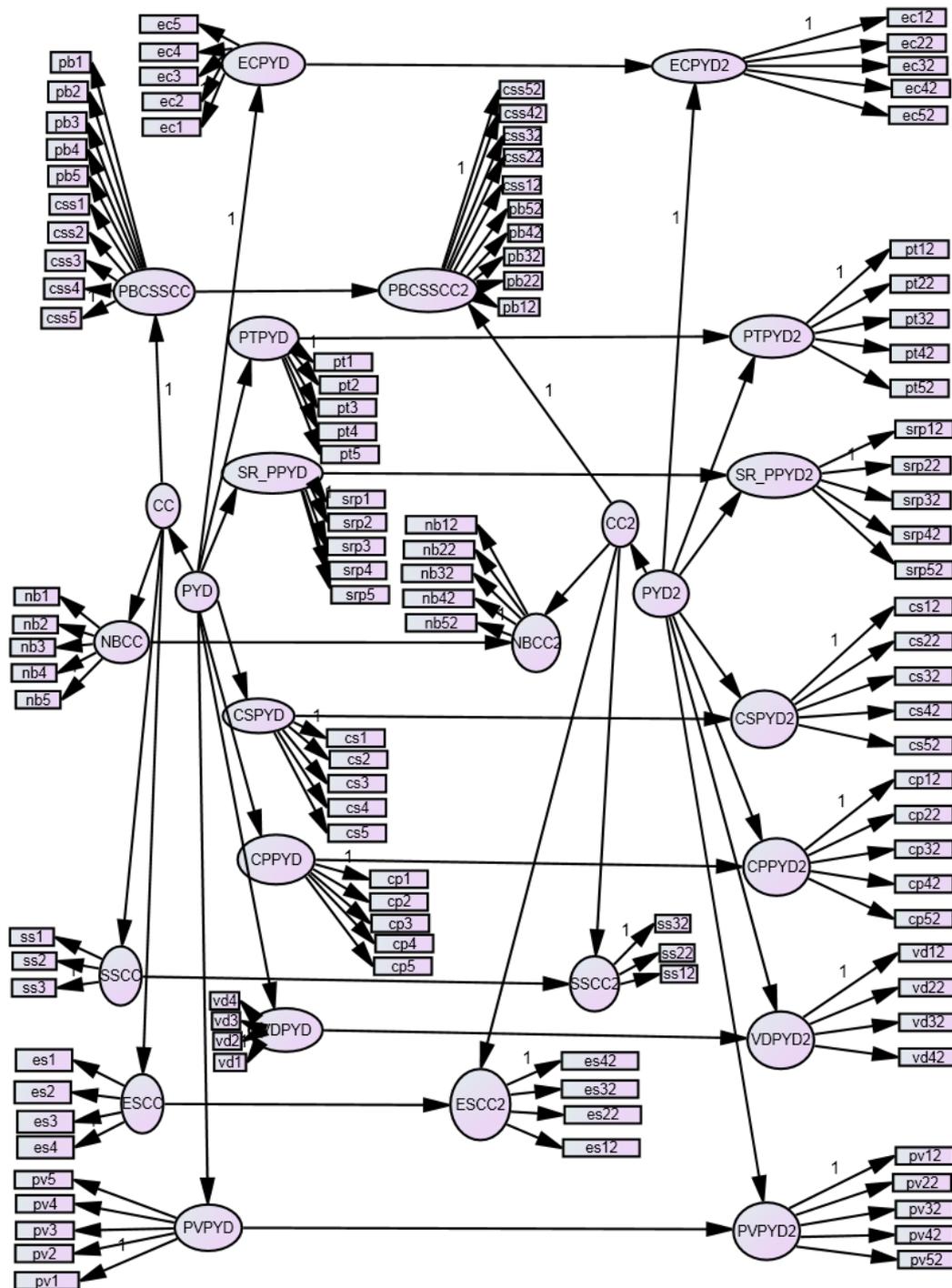
*Note.* ECPYD= empathic concern (Time 1); PTPYD=perspective taking (Time 1); SRPPYD=social rules/politeness (Time 1); CSPYD=connection to school (Time 1); CPPYD= connection to peers (Time 1); VDPYD=values diversity (Time 1); PVPYD=personal values (Time 1); ESCC= enjoyment of school (Time 1); SSCC=safety at school (Time 1); NBCC=negative behaviour (Time 1); PBCSSCC=positive behaviour and classroom supportiveness (Time 1); ECPYD2=empathic concern (Time 2); PTPYD2=perspective taking (Time 2); SRPPYD2=social rules/politeness (Time 2); CSPYD2= connection to school (Time 2); CPPYD2=connection to peers (Time 2); VDPYD2=values diversity (Time 2); PVPYD2=personal values (Time 2); ESCC2= enjoyment of school (Time 2); SSCC2 =safety at school (Time 2); NBCC2=negative behaviour (Time 2); PBCSSCC2=positive behaviour and classroom supportiveness (Time 2).

\*\*\*=  $p < .001$ .

The magnitude of the autoregression effect sizes was large ( $f^2 = .35$  to  $.61$ ). However, to test for other factors that may have influenced scores at Time 2, Model 8 added cross-lags from the PYD factors to the non-identical CC factors. The cross-lag paths test the effect of Time 1 scores for PYD on Time 2 CC variables. Model 8A tested PYD1 factors cross-lagged onto CC2 factors, while Model 8B tested CC1 factors cross-lagged onto PYD2 factors. In addition, trimmed models that removed statistically non-significant paths ( $p \geq .05$ ) were also tested (i.e., Models 8-2 and 8A-2). The fit statistics and estimate parameters for Models 8, 8-2, 8A, and 8A-2 are reported in Table 32.

All four versions of Model 8 were tested for group invariance. Model 8A-2 was invariant for groups across the configural, metric, and scalar levels. Models 8 and 8A were not invariant, for the experienced Kiwi Can group level ( $\Delta\text{CFI} > .01$ ). On this basis, Models 8, 8-2, and 8A were deemed unacceptable. This left two plausible SEM models – the full AR model (Model 7) and the trimmed cross-lagged CC to PYD model (Model 8A-2). The two models were compared for relative fit to the data using the Wilson-Hilferty (1931) and AIC procedures. The Wilson-Hilferty indicated that Model 7 had significantly better fit than Model 8A-2 ( $z_{\text{diff}\chi^2} = 20.23$ ,  $df = 4.90$ ,  $z = 15.33$ ,  $p < .001$ ), and the AIC weight ( $\Delta I = 1.31$ ) also supported Model 7. Therefore, the full AR model without cross-lags was retained as the final representation of structural change across time (see Figure 13).

**Figure 13.** Final Autoregressive Structural Equation Model for PYD and CC



*Figure 13.* A copy of the final autoregressive PYD-CC SEM. ECPYD=empathic concern; PTPYD=perspective taking; SRPPYD=social rules/politeness; CSPYD=connection to school; CPPYD=connection to peers; VDPYD=values diversity; PVPYD= personal values; NBCC= negative behaviour; SSCC=safety in school; ESCC= enjoyment of school; PBCSSCC=positive behaviour and classroom supportiveness.

## 6.11 Summary

These analyses indicated that individual social developmental factors (PYD) predicted variability in classroom climate level factors within each cross-section of time, but not across time. Across time, scores for each factor at Time 2 were most strongly predicted by the starting score for the same factor. This finding was surprising given that positive youth development scholars have previously reported that the accumulation of assets, including a positive school environment, are associated with superior individual positive developmental outcomes (Benson & Leffert, 2001; Scales, 1999). In this case, a positive classroom climate did not serve as an asset or resource that contributed to superior prosocial development over time. It was also surprising because the results do support the widespread belief (Bensen et al., 1998) that individual connection, caring/compassion, competence, and character contribute to a highly prosocial classroom environment. Instead, the results from this study suggest that the final AR model should serve as the basis for all future analysis.

## 6.12 Results - Mean Differences in PYD and CC Subscales

In this section the baseline means for individual PYD and CC factors are presented.

### 6.12.1 Baseline means – PYD subscales.

The baseline mean scores for PYD subscales ranged from  $M = 4.59$ ,  $SD = .88$  (Perspective Taking) to  $M = 5.31$ ,  $SD = .68$  (Social Rules/Politeness), with an overall mean of  $M = 4.97$ ,  $SD = .82$  (see Appendix T for a complete table of PYD subscales means at baseline). When means were compared within groups using MANOVAs, Pillai's trace results indicated that there was a significant effect of school condition,  $V = .03$ ,  $F_{(14, 2422)} = 2.39$ ,  $p = .003$ , sex,  $V = .08$ ,  $F_{(7, 1211)} = 13.97$ ,  $p < .001$ , year level,  $V = .08$ ,  $F_{(14, 2418)} = 6.81$ ,  $p < .001$ , and ethnicity,  $V = .14$ ,  $F_{(28, 4824)} = 6.13$ ,  $p < .001$  on the mean scores.

According to the Games-Howell post hoc test results and pairwise comparisons (see Table 43), there were very small statistically significant differences in baseline mean scores for two of the seven PYD subscales as a function of school condition. They were reported for Perspective Taking,  $F_{(2, 1216)} = 6.24$ ,  $p = .002$  and Social

Rules/Politeness,  $F_{(2, 1216)} = 5.46, p = .004$ . In these two subscales the mean scores were statistically significantly lower for the experienced Kiwi Can school children than for the control and/or new school children. School condition also had a statistically significant but trivial effect on baseline scores for Empathic Concern,  $F_{(2, 1216)} = 3.85, p = .022$ , Connection to School,  $F_{(2, 1216)} = 3.64, p = .027$ , Connection to Peers,  $F_{(2, 1216)} = 5.01, p = .007$ , and Personal Values,  $F_{(2, 1216)} = 4.68, p = .009$ . There was no statistically significant effect of school condition on baseline scores for Values Diversity,  $F_{(2, 1216)} = 1.21, p = .299$ . These results were not expected. In fact, they are the opposite of the initial hypothesis that experienced Kiwi Can students would report higher mean subscale scores for PYD than either new or control school students. However, it must be noted that these results reflect very small to trivial differences and that the baseline mean scores for PYD and CC means were not deemed to be meaningfully different between experienced Kiwi Can, new Kiwi Can, and control school students.

Table 43

*Games-Howell Post Hoc Test Results (MANOVA) for Effect of School Condition on PYD Subscale Scores at Baseline*

Dependent Variable	$F(2, 1216)$	Effect size Differences $\eta^2$	Effect Size Differences Cohen's $f^2$
Perspective Taking	6.24**	.010	.01
Social Rules / Politeness	5.46**	.009	.01
Connection to Peers	5.01**	.008	< .01
Personal Values	4.68**	.008	< .01
Empathic Concern	3.85*	.006	< .01
Connection to School	3.64*	.006	< .01
Values Diversity	1.21 <sup>ns</sup>	.002	< .01

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ , \* =  $p < .05$ , ns = not statistically significant.

Baseline means were also statistically significantly lower for boys than girls across all seven PYD subscales ( $p < .001$ ). The effect of sex ranged from trivial to small  $.01 < f^2 > .06$ . They were largest for measures of caring and compassion (Empathic Concern, Perspective Taking), social competence (Social Rules/Politeness), and character (Personal Values). When differences were compared as a function of year level, children in the oldest age cohort (Year 8) reported statistically significantly lower mean scores than their younger peers (Year 6 or Year 7) for six of the seven PYD subscales. There were no significant differences in mean scores for the

Connection to Peers subscale as a function of age. Effect sizes were very small for the Social Rules/Politeness, Empathic Concern, Perspective Taking, and Values Diversity subscales ( $f^2 = .01$  to  $.02$ ), but grew in magnitude slightly for the Personal Values ( $f^2 = .03$ ) and Connection to School subscale ( $f^2 = .04$ ). Finally, the results of the MANOVAs for ethnicity, showed that Māori students had statistically significantly lower mean scores than Pakeha students for the Empathic Concern, Perspective Taking, Social Rules/Politeness, and Personal Values subscales, and lower mean scores than Pasifika for the Values Diversity subscale. There was no effect of ethnicity on the Connection to Peers subscale (see Table 44). The size of the effect was trivial for Connection to School and Perspective Taking ( $f^2 = .01$ ), and small for Values Diversity, Social Rules/Politeness, Personal Values, and Empathic Concern ( $.02 < f^2 > .05$ ).

Table 44

*Games-Howell Post Hoc Test Results (MANOVA) for Effect of Sex, Year Level, and Ethnicity on PYD Subscale Scores at Baseline*

Dependent Variable	<i>df</i>	<i>F(p)</i>	Effect size differences $\eta^2$	Effect Size Differences Cohen's $f^2$
<b>Sex</b>				
Empathic Concern	1, 1217	72.57***	.056	.06
Social Rules/Politeness	1, 1217	62.04***	.049	.05
Perspective Taking	1, 1217	51.78***	.041	.04
Personal Values	1, 1217	48.02***	.038	.04
Values Diversity	1, 1217	39.86***	.032	.03
Connection to School	1, 1217	13.82***	.011	.01
Connection to Peers	1, 1217	11.62**	.009	.01
<b>Year Level</b>				
Connection to School	2, 1214	26.71***	.042	.04
Personal Values	2, 1214	15.34***	.025	.03
Values Diversity	2, 1214	14.77***	.024	.02
Perspective Taking	2, 1214	9.54***	.015	.02
Empathic Concern	2, 1214	9.36***	.015	.02
Social Rules/Politeness	2, 1214	8.55***	.014	.01
Connection to Peers	2, 1214	1.86 <sup>ns</sup>	.003	< .01
<b>Ethnicity</b>				
Values Diversity	4, 1209	17.28***	.054	.05
Social Rules/Politeness	4, 1209	8.32***	.027	.03
Personal Values	4, 1209	6.34***	.021	.02
Empathic Concern	4, 1209	5.87***	.019	.02
Connection to School	4, 1209	4.09**	.013	.01
Perspective Taking	4, 1209	3.53**	.012	.01
Connection to Peers	4, 1209	2.10 <sup>ns</sup>	.007	< .01

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ , \* =  $p < .05$ ; ns=not statistically significant.

**6.12.2 Baseline means – CC subscales.** The baseline mean scores for classroom climate CC, ranged from  $M = 3.35$ ,  $SD = 1.11$  to  $M = 4.94$ ,  $SD = .90$ , with an overall mean of  $M = 4.50$ ,  $SD = .97$  (see Appendix U for a complete table of baseline means for all CC subscales). When CC subscale means were compared within groups, the Pillai's trace statistic indicated that there was a significant but very small main effect of school condition,  $V = .04$ ,  $F_{(8, 2428)} = 5.49$ ,  $p < .001$ , sex,  $V = .02$ ,  $F_{(4, 1214)} = 5.27$ ,  $p < .001$ , year level,  $V = .07$ ,  $F_{(8, 2424)} = 11.45$ ,  $p < .001$ , and ethnicity,  $V = .08$ ,  $F_{(16, 4836)} = 6.15$ ,  $p < .001$ . For school condition and sex, the Pillai's trace values,  $V$ , were  $\leq .10$  which suggests that the effect was weak and practically meaningless. On this basis, the decision was made to dispense with any further analysis of these two independent effects. Year level and ethnicity, on the other hand, had Pillai's trace statistics approaching  $V = .10$ . Differences in CC subscale means were explored using MANOVAs (see Table 45). According to the output, Year 8 students reported statistically significantly lower mean scores than Year 6 or 7 students for Positive Behaviour and Classroom Supportiveness,  $F_{(2, 1216)} = 11.38$ ,  $p < .001$ ,  $f^2 = .02$ , Safety at School,  $F_{(2, 1216)} = 17.75$ ,  $p < .001$ ,  $f^2 = .03$ , and Enjoyment of School,  $F_{(2, 1216)} = 33.90$ ,  $p < .001$ ,  $f^2 = .05$ . The size of the effect was small for each subscale. For ethnicity, baseline Positive Behaviour and Classroom Supportiveness,  $F_{(2, 1216)} = 5.83$ ,  $p < .001$  and Enjoyment of School,  $F_{(2, 1216)} = 7.92$ ,  $p < .001$  means were significantly lower for Māori than Pasifika students. Subscale means for Negative Behaviour,  $F_{(2, 1216)} = 12.06$ ,  $p < .001$  were also statistically significantly higher at baseline for Māori than Pakeha students. All effect sizes were small ( $f^2 < .15$ ).

Table 45

*Games-Howell Post Hoc Test Results (MANOVA) for Effect of Year Level, and Ethnicity on CC Subscale Scores at Baseline*

Dependent Variable	<i>df</i>	<i>F(p)</i>	Effect Size Differences $\eta^2$	Effect Size Differences Cohen's $f^2$
Year Level				
Enjoyment of School	2, 1216	33.90***	.053	.05
Safety at School	2, 1216	17.75***	.028	.03
Positive Behaviour and Classroom Supportiveness	2, 1216	11.38***	.018	.02
Negative Behaviour	2, 1216	5.86 <sup>ns</sup>	.010	.01
Ethnicity				
Negative Behaviour	4, 1214	12.06***	.038	.04
Enjoyment of School	4, 1214	7.92***	.026	.03
Positive Behaviour and	4, 1214	5.83***	.019	.02

Classroom Supportiveness				
Safety at School	4, 1214	2.77*	.009	.01

\*\*\* =  $p < .001$ , \*\* =  $p < .01$ , \* =  $p < .05$ ; ns = not significant.

### 6.12.3 Discussion of baseline PYD and CC subscale means.

The average baseline mean scores for the PYD subscales were high ( $M > 4.00$ ), indicating that children felt they were caring/compassionate, socially competent, connected to their school and peers, valued others, and engaged in prosocial behaviours at the beginning of the school year. They also reportedly felt safe in their classrooms and enjoyed being a student in their respective schools. When queried about the behaviour of their classmates, high mean CC subscale scores ( $M > 3.55$ ) indicated that individuals felt that their peers behaved positively in the classroom, and showed care, consideration, and support for one another. When PYD subscale mean scores were compared between groups, there were trivial or very small effects on subscale scores for school condition and ethnicity. Effect sizes grew slightly as a function of sex, though they remained small ( $f^2 < .15$ ) for all subscales.

Year level also slightly affected three of the four CC subscale means – Positive Behaviour and Classroom and School Supportiveness, Enjoyment of School, and Safety at School, at baseline. The effects were all very small. Small effects were also reported for Negative Behaviour, Enjoyment of School, and Positive Behaviour and Classroom Supportiveness as a function of ethnicity.

Given that the mean differences in all subscales were very small and/or trivial as a function of school condition, year level, sex, and ethnicity it was assumed that these variables had little to no significant effect on children’s social attitudes and behaviours or classroom climate at baseline. They were therefore not expected to confound the results of the analysis of change in mean subscale scores across time. For this analysis, changes in mean subscale scores were explored as a function of school condition and participant characteristics (sex, year level, and ethnicity).

Mean scores for all PYD and CC subscales decreased from baseline (pre-test) to the end of the school year (post-test). The magnitude of change varied both between and within groups (see Appendices V and W for a complete record of the mean PYD and CC subscale change scores). The change in subscale scores also varied as a function

of school condition and participant characteristics. Results of the one-way, repeated measures analysis of variance (RM-ANOVA) tests are presented below. Unless otherwise stated the sphericity assumed estimate was used for all data analyses in this section.

#### 6.12.4 Change in PYD and CC subscale means over time.

##### 6.12.4.1 PYD by school condition.

Contrary to expectations and initial hypotheses, results of the one-way RM-ANOVAs indicated that there were extremely small, significant effects of school condition on Empathic Concern,  $V = .01$ ,  $F_{(2, 1216)} = 5.93$ ,  $p = .003$  and Values Diversity,  $V = .01$ ,  $F_{(2, 1216)} = 7.82$ ,  $p < .001$ . Pairwise comparison output showed that experienced school children reported a larger decrease in Empathic Concern than both new and control school children ( $p < .01$ ). Experienced Kiwi Can school children also reported larger decreases in Values Diversity mean scores over time than new school or control school children ( $p < .05$ ). Effects of school condition were also significant but trivial for the change in Connection to School subscale scores over time,  $V = .007$ ,  $F_{(2, 1216)} = 4.12$ ,  $p = .016$ . There were no statistically significant effects of school condition on Perspective Taking,  $V = .001$ ,  $F_{(2, 1216)} = .733$ ,  $p = .481$ , Social Rules/Politeness,  $V = .003$ ,  $F_{(2, 1216)} = 1.71$ ,  $p = .181$ , Connection to Peers,  $V = .001$ ,  $F_{(2, 1216)} = .63$ ,  $p = .534$ , and Personal Values,  $V = .003$ ,  $F_{(2, 1216)} = 1.74$ ,  $p = .176$ . See Table 46 for a summary of the one-way RM-ANOVA output.

Table 46

*One-Way Repeated Measures Analysis of Variance Results and Effect Sizes for Pre-Post PYD Subscale Means as a Function of School Condition*

Dependent Variable	$F(2, 1216)$	$p$	Effect size differences $\eta^2$	Effect size differences Cohen's $d$
Values Diversity	7.82	***	.013	.01
Empathic Concern	5.93	**	.010	.01
Connection to School	4.12	*	.007	.01
Personal Values	1.74	.176	.003	<.01
Social Rules/ Politeness	1.71	.181	.003	<.01
Perspective Taking	.73	.481	.001	<.01
Connection to Peers	.63	.534	.001	<.01

\*\*\* =  $p < .001$ ; \*\* =  $p < .01$ ; \* =  $p < .05$ .

#### **6.12.4.2 PYD by participant characteristics.**

Neither sex nor year level had a statistically significant effect on the change in PYD subscale means ( $p \geq .05$ ). Ethnicity, on the other hand, had a significant but trivial main effect on three PYD subscales – Perspective Taking,  $F_{(4, 1209)} = 2.74, p = .028, f^2 < .01$ , Social Rules/Politeness,  $F_{(4, 1209)} = 2.73, p = .028, f^2 < .01$ , and Values Diversity,  $F_{(4, 1209)} = 3.54, p = .007, f^2 < .01$ . Given that the effects were either trivial in size and/or insignificant, these results do not bear further discussion or reporting.

#### **6.12.4.3 CC by school condition, sex, year level, and ethnicity.**

Results of the one-way RM-ANOVAs indicated that school condition had a statistically significant but trivial effect on the change in two subscale means over time – Negative Behaviour,  $F_{(2, 1216)} = 2.96, p = .052, f^2 < .01$  and Safety at School,  $F_{(2, 1216)} = 3.19, p = .037, f^2 < .01$ . Results also showed that the other two classroom climate subscales, Positive Behaviour and Classroom Supportiveness,  $F_{(2, 1216)} = .10, p = .905$  and Enjoyment of School,  $F_{(2, 1216)} = 2.77, p = .063$  were not significantly affected by school condition. Neither sex nor year level had a statistically significant effect on the change in CC subscale means ( $p > .05$ ). Ethnicity, on the other hand, had three statistically significant but trivial main effects on Positive Behaviour and Classroom and School Supportiveness,  $F_{(4, 1209)} = 4.17, p = .002, f^2 < .01$ , Negative Behaviour,  $F_{(4, 1209)} = 3.06, p = .016, f^2 < .01$ , and Enjoyment of School,  $F_{(4, 1209)} = 2.72, p = .028, f^2 < .01$ . For Positive Behaviour and Classroom and School Supportiveness and Enjoyment of School, students who identified as Māori reported significantly larger decreases in mean subscale scores than Pakeha or Asian students. For the Negative Behaviour subscale Māori students reported significantly smaller decreases than their Pakeha peers over time.

#### **6.12.5 Discussion of change in PYD and CC subscale mean results.**

The results of the outcome evaluation indicate that there were few statistically significant differences in measures of PYD or CC for the Year 6-8 participants. Further, when differences were identified in mean change scores as a function of school condition, the magnitude of difference was small and/or trivial in each case ( $.01 < f^2 > .05$ ). This suggests that there were no practically significant differences in the

self-reported incidence of prosocial attitudes and/or behaviours between children across experienced Kiwi Can, new Kiwi Can, and/or control schools.

While it was expected that the participant characteristics would not affect the change in mean PYD and CC scores (and they did not), it was expected that school condition would produce significant differences in mean results. In fact, it was hypothesised that experienced and new school children would report gains in both PYD and CC subscale means across time. Further, it was expected that new Kiwi Can school children would report larger changes in mean scores for individual social development and classroom climate than those of either experienced or control school children. Finally, it was expected that the mean scores for PYD or CC subscales would remain the same for control school children over time. None of these three hypotheses were substantiated.

Firstly, instead of reporting increases in mean PYD and CC subscale scores over time, experienced and new Kiwi Can school children reported decreases in all measures of positive, prosocial attitudes, and behaviours. They also reported decreases in the incidence of positive, prosocial behaviour amongst their classmates and lower levels of enjoyment and safety within their schooling environment. The decreases were small but statistically significant for measures of caring (Empathic Concern), character (Values Diversity), and connection to school (Connection to School). In each case, the mean change was significantly larger for experienced schools than either new and/or control schools. Secondly, new Kiwi Can school children did not report larger changes in PYD and/or CC mean scores over time. In fact, new Kiwi Can school children reported mean decreases that were not dissimilar from control school children across all eleven subscales, and not significantly different in a practical and meaningful way from experienced school children across eight of the eleven subscales. For the three subscales in which small, significant differences did arise (Empathic Concern, Values Diversity, and Connection to School), the new Kiwi Can schools reported smaller mean differences than the experienced Kiwi Can schools. Finally, mean scores for PYD and CC subscales did not remain the same for the control group. Instead, all measures of social attitudes and behaviours decreased from pre-test to post-test for this group.

The results from this outcome evaluation of Kiwi Can do not suggest that participation in a school-based, positive social development programme contributes to the development of the prosocial attitudes at the individual level that enable children to demonstrate caring and compassion for others, behave in socially competent ways, feel connected to their school or peers, demonstrate acceptance of diversity, or value difference in their social world. Nor do they suggest that participation in a school-based positive youth development programme improves the social climate of the school classroom. Instead, in some cases, children reported less prosocial and more antisocial behaviour amongst their peers after a year of Kiwi Can. Many of these same children reported feeling less safe in their school classroom and on their school grounds. They also reported that they didn't enjoy attending school in late 2013 as much as they had at the beginning of the year.

These results were not expected. Thus, additional analyses were undertaken to explore what, if any, moderating, mediating, and/or confounding factor(s) may have affected the current results. Firstly, the interaction between individual demographic characteristics (sex, year level, and ethnicity) and school condition was investigated using a series of factorial RM-ANOVAs (one for each PYD and CC subscale). Four sets of factorial RM-ANOVAs were run for – (i) school condition by ethnicity; (ii) school condition by year level; and (iii) school condition by sex. In each case the factors were treated as independent variables and the subscales as dependent variables. At the end of the process, it was clear that there were no significant differences ( $p \geq .05$ ) in the results when the interaction effects of school condition by ethnicity, year level, and/or sex were considered. Therefore, the decision was made to explore factors at the school level that might have an impact on student social attitudes and behaviours. Transience was explored because the participants all came from low-decile schools and recent reports had identified high rates of transience in schools from less affluent areas of the country (Child Poverty Action Group, 2014b; McLeod, 2014). Also there had been a high rate of student turn over amongst children in this study with hundreds of children completing outcome evaluation surveys at only one time point and not the other. This indicated that they may have either moved away from, or moved into, the participating school over the course of the 2013 school year. An operational definition for transience was created, and this was used to compute transience scores for each participating school and to categorise

schools into new groups by degree of transience rather than by programme participation. Differences in mean scores for PYD and CC subscales were then compared at baseline and across time as a function of transience.

### **6.13 Exploring Changes in PYD and CC Subscale Means as a Function of Transience**

In this section the results from an exploratory study on the effect of transience on self-reported prosocial attitudes and behaviours in Year 6-8 students from experienced Kiwi Can, new Kiwi Can, and control schools are presented. The section commences with a short discussion of how transience was defined and operationalised in this study. It is followed by the results of a series of MANOVA tests exploring the effect of transience condition on PYD and CC subscale means at baseline. The effect of transience on change in PYD and CC subscale means is then explored, using one-way RM-ANOVA tests. The section concludes with the results of an exploratory study in which the interactive effects of transience and school condition were investigated cross-sectionally, using MANOVAs with two-fixed factors, and across time using one-way RM-ANOVAs with one fixed, between subjects factor (School x Transience). The disaggregated, RM-ANOVA approach was selected for this analysis rather than the hierarchical or multilevel linear modeling (HLM) approach for two reasons. Firstly, the number of groups at level 3 was so small (< 20) that there were concerns about the power of HLM to adequately detect significant differences in effect at the group level. Secondly, there was little to no evidence of substantial between-groups variance as evidenced by the small (i.e. < .10 or 10%) intra-class correlation (ICC) co-efficients computed for each PYD and each CC subscale (Woltman, Feldstain, MacKay, & Rocchi, 2012). The ICCs were computed using the Mixed Linear Model feature in SPSS. The final ICC values were as follows: EC (5.40%), PT (< .01%), SRP (.70%), CS (8.87%), Co (.92%), VD (< .01%), PV (< .01%), PBCSS (3.85%), NB (1.63%), SS (1.29%), and ES (3.17%).

### 6.13.1 Defining transience.

The notion of transience in public school settings can be operationalised and measured both at the individual level and at the school level. It is a measure of change that reflects the degree to which children move in and/or out of a school over a period of time. Typically, transience is measured cross-sectionally and reflects change in the school roll over the course of one academic year. In New Zealand, school-level transience has been explored in the primary school setting and measured according to a formula generated by Anne Johnson (2002), in a study on transience in South Auckland schools. According to the formula, transience is equal to the difference in enrolments from the beginning of the year to the end of the school year, divided by the average school roll. While Johnson's work served as an important reference for this thesis it was not appropriate to adopt the operational definition of transience used by Johnson in this study. This was because information about the change in student rolls for the schools participating in this study was not available in the public domain, and was therefore inaccessible to the researcher. Further, there was an interest in understanding transience as it pertained specifically to the collection of survey data in this study. Therefore, a modified mathematical formula for computing rates of school-level transience was created. It was based on the assumption that children who completed both Kiwi Can outcome evaluation surveys (pre-post) were non-transient (i.e., had not moved out of the school) while those who completed surveys at only one time point may be transient. It was acknowledged, however, that these potentially transient children may simply have been absent on the day of data collection and that a mechanism must be introduced into the transience computation to account, or correct, for this possibility. In the transience formula, this correction value was termed the *Legit score*. It incorporated school roll data at Time 1 and compared the number of children who had completed the survey with the number of children who attended the school. In this thesis, the school roll data were obtained from the TKI website (<http://www.tki.org.nz/Schools>). A legit score of .90 indicated that 90% of the students on the school roll had completed the social health survey. In this study the legit values ranged from .86 to 1.24. Legit values greater than 1.0 were not expected but may reflect the fact that the school roll data obtained for the study may not have been entirely up-to-date. However, it was the most

accurate source of data available. The school-level transience formula adopted for this thesis is:

$$\text{School-level transience (\%)} = (Nap+Npa)/(NT1+Nap+Npa)(Legit\ score)$$

$Nap$  = Number of students who did not complete a survey at Time 1 ('absent') but completed a survey at Time 2 ('present')

$Npa$  = Number of students who were present at Time 1 and absent at Time 2

$NT1$  = Number of students who completed the survey at Time 1

$Nenroll$  = Number of students enrolled in the year groups surveyed at the school

$Legit\ score$  =  $NT1/Nenroll$

The  $Nap$ ,  $Npa$  and  $NT1$  data was generated by exploring the original data sets from Time 1 (T1 = start of the year) and Time 2 (T2 = end of the year). The  $Nenroll$  is an estimate of the number of children in each school potentially enrolled in the study. The total roll for each of the participating schools was divided by the number of year groups taught in the school, and then multiplied by the number of years surveyed in this study. For example:

School #1 was a full primary school with eight year groups and a roll of 103 students, but only Years 6-8 were surveyed. Hence:

$$Nenroll = 103 *(3/8) = 38.63$$

On the basis of this formula, the transience rate was calculated for each of the 24 participating schools (see Table 47). Percentiles were then calculated to create three test conditions for this study – *low transient schools*, *middle transient schools*, and *high transient schools*. Low transient schools had transience means in the lowest 25<sup>th</sup> percentile of total scores, while high transient schools had mean values at or above the 75<sup>th</sup> percentile. The average rate of transience for all participating schools in this study was approximately 24%.

Table 47

*The Nap, Npa, NTI, Nenroll, Legit and Transience Scores for the Participating Schools*

School	NT1	Nenroll	Legit score	Npa	Nap	Transience Rate (%)
1.	32	34.33	.93	7	3	25.50
2.	6	6.60	.91	3	2	50.00
3.	64	69.33	.92	23	1	29.50
4.	43	46.50	.93	8	15	37.70
5.	24	23.14	.96	6	0	20.70
6.	30	35.33	.92	3	9	31.30
7.	79	86.00	.92	20	17	34.70
8.	257	275.50	.93	32	56	27.30
9.	226	194.50	1.16	19	32	15.80
10.	29	28.44	1.02	13	2	33.40
11.	29	31.86	.91	11	0	30.20
12.	19	19.14	.99	3	3	24.20
13.	277	305.00	.91	42	40	25.20
14.	7	8.14	.86	1	5	53.70
15.	20	19.00	1.05	9	6	40.80
16.	28	27.86	1.01	4	3	19.80
17.	30	31.71	.95	5	0	15.10
18.	27	25.71	1.05	4	2	17.30
19.	24	19.33	1.24	3	1	11.50
20.	282	301.14	.94	20	28	15.50
21.	89	97.17	.92	11	3	14.80
22.	27	29.29	.92	4	0	14.00
23.	5	5.14	.97	0	0	0.00
24.	62	67.71	.92	0	5	8.10
<b>Average</b>	<i>M</i> = 24.84	<i>SD</i> = 12.87				
<b>25<sup>th</sup></b>	<b>percentile</b>	15.20				
<b>75<sup>th</sup></b>	<b>percentile</b>	32.88				

Note. *NTI* = number of students who completed the survey at Time 1; *Nenroll* = Number of students enrolled in the year groups surveyed at the school; *Legit score* = *NTI*/*Nenroll*; *Nap*= Number of students who did not complete a survey at Time 1 but completed a survey at Time 2; *Npa*= Number of students who were present at Time 1 and absent at Time 2.

On the basis of this output, a total of six schools were categorised as low transience ( $n = 178$  student participants), 13 schools were categorised as middle transience ( $n = 928$  student participants), and five schools were categorised as high transience ( $n = 113$  student participants). All schools in the low transience group ( $n = 6$ ) were control schools. In the middle transient group there were three control, five new, and four experienced schools. In the high transient group there were three experienced and two new Kiwi Can schools.

Dummy variables were generated to represent the transience condition and entered into the final data set (1 = *low transience*, to 3 = *high transience*). Mean scores for PYD and CC subscales were then compared at baseline. It should be noted that when the results of the Box's test of equality of covariance matrices were significant at  $p <$

.01 - .05, Q-Q plots, and covariance matrices were computed and manually inspected for evidence of non-normality and a lack heterogeneity in covariance structures. Unless otherwise stated, the Q-Q plots appeared normal and the decision was made to assume homogeneity.

Given the vastly different sample sizes between groups, the Pillai-Bartlett trace ( $V$ ) and Games-Howell post hoc tests were employed for all MANOVAs in this section. For the analysis of difference in mean scores across time, one-way RM-ANOVAs and Factorial RM-ANOVAs were employed. One-way RM-ANOVAs were used to explore the interactive effect of transience and school condition on PYD and CC means. Unless otherwise stated, the Sphericity Assumed estimate and the Games-Howell post hoc test were used in all MANOVA and one-way RM-ANOVA analyses.

#### **6.14 Results - Baseline Means (Transience)**

In this exploratory study, the means for the PYD and CC subscales were investigated at baseline as a function of transience. This information was gathered in order to establish what, if any, differences existed in subscale means for children from experienced Kiwi Can, new Kiwi Can, and control schools. Differences in baseline means were first computed for PYD subscales and then for CC subscales. The results of these analyses are below.

##### **6.14.1 Baseline means – PYD subscales.**

The baseline means for all PYD subscales ( $n = 7$ ) varied as a function of transience (see Table 48 for the baseline means, standard deviations, and confidence intervals for subscale by transience condition). Means for two of the seven PYD subscales differed significantly as a function of transience, Pillai's  $V = .05$ ,  $F_{(12, 2422)} = 4.20$ ,  $p = .001$ . According to the post hoc output, there were statistically significant mean differences in the Connection to School,  $F_{(2,126)} = 10.84$ ,  $p \leq .001$ ,  $f^2 < .01$  and Connection to Peers,  $F_{(2, 126)} = 6.82$ ,  $p = .001$ ,  $f^2 = .19$  subscales. While the size of the effect of transience on Connection to School was trivial, the effect for Connection to Peers was medium. In this case, students from low transient schools reported statistically significantly higher mean scores for Connection to Peers at baseline than

did students from high transient schools ( $p < .05$ ). There were no other statistically significant differences identified in PYD subscales as a function of transience.

Table 48

*Means, Standard Deviations, and Confidence Intervals for PYD Subscales as a Function of Transience*

Dependent Variable	Transience Condition	<i>n</i>	<i>M</i>	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
Empathic Concern	Low	178	5.08	.82	4.95	5.21
	Middle	928	5.08	.88	5.03	5.14
	High	113	5.15	.83	4.99	5.31
Perspective Taking	Low	178	4.67	.83	4.54	4.80
	Middle	928	4.58	.89	4.52	4.63
	High	113	4.58	.88	4.42	4.75
Social Rules/Politeness	Low	178	5.42	.64	5.32	5.52
	Middle	928	5.29	.68	5.25	5.33
	High	113	5.31	.70	5.18	5.43
Connection to School	Low	178	5.28	.62	5.17	5.40
	Middle	928	5.00	.82	4.94	5.05
	High	113	5.14	.73	4.99	5.28
Connection to Peers	Low	178	5.02	.77	4.88	5.16
	Middle	928	4.81	.97	4.75	4.87
	High	113	4.61	.93	4.44	4.79
Values Diversity	Low	178	4.85	.84	4.73	4.98
	Middle	928	4.81	.84	4.76	4.87
	High	113	5.00	.84	4.84	5.15
Personal Values	Low	178	5.08	.77	4.96	5.20
	Middle	928	4.98	.80	4.93	5.04
	High	113	5.02	.89	4.87	5.16

*Note.* CI= confidence interval, LL = lower limit, UL = upper limit.

#### 6.14.2 Baseline means - CC subscales.

The baseline means for CC also varied as a function of transience (see Table 49 for the means, standard deviations, and confidence intervals for baseline CC means by transience level). Results of the Pillai's trace indicated that there was a significant difference in mean CC baseline scores as a function of transience, Pillai's  $V = .02$ ,  $F_{(8, 2428)} = 3.34$ ,  $p = .001$ . The post hoc test output showed that there were small ( $f^2 = .02$ ) statistically significant differences in the Safety at School,  $F_{(2, 1216)} = 4.35$ ,  $p = .013$  and Enjoyment of School subscales,  $F_{(2, 1216)} = 6.73$ ,  $p = .001$ . According to the pairwise comparison output, children in the low transient schools reported statistically significantly higher mean scores for both subscales than children in middle transient school conditions. There were no statistically significant differences in self-reported

means for the Positive Behaviour and Classroom Supportiveness, or Negative Behaviour subscales.

Table 49

*Comparison of Means for CC Factors at Time 1 Between Students from Low, Middle, and High Transience Schools*

Dependent Variable	Transience Condition	n	M	SD	95%CI	
					LL	UL
Positive Behaviour and Classroom and School Supportiveness	Low	178	4.77	.76	4.65	4.89
	Middle	928	4.71	.82	4.65	4.76
	High	113	4.79	.79	4.64	4.94
Negative Behaviour	Low	178	3.58	1.06	3.42	3.75
	Middle	928	3.64	1.11	3.57	3.71
	High	113	3.81	1.12	3.60	4.01
Safety at School	Low	178	5.10	.81	4.97	5.24
	Middle	928	4.89	.92	4.83	4.94
	High	113	4.94	.94	4.77	5.11
Enjoyment at School	Low	178	5.20	.82	5.04	5.36
	Middle	928	4.89	1.11	4.82	4.96
	High	113	5.02	1.07	4.82	5.21

Note. CI= confidence interval, LL = lower limit, UL = upper limit.

### 6.14.3 Discussion of baseline means.

Overall, there was little statistically significant effect for transience on baseline means. The data showed that where statistically significant differences arose in the two PYD (Connection to School, Connection to Peers) and in the two CC subscales (Enjoyment of School, Safety at School) the size of the effect was very small and/or trivial. In all four cases, children in low transient schools reported statistically significantly higher mean scores than children in middle and/or high transient schools. Overall, these results suggest that the majority of children from all schools had similar social attitudes and behaviours at the start of the 2013 school year. The similar baseline means were, therefore, not expected to confound the results of the analysis of change in mean subscale scores across time.

### 6.15 Results - Change in PYD and CC Subscale Means Over Time (Transience)

The one-way repeated measures analysis of variance procedure was then used to explore changes in social attitudes and behaviour over time as a function of transience

level. The mean scores for all PYD and CC subscales decreased over time as a function of transience for students in high and middle transient schools, while within the low transience group, mean scores for the Values Diversity, Personal Values, and Safety at School subscales increased by a small amount. See Table 50 for a copy of the mean change scores, standard deviations, and confidence intervals for PYD and CC subscale means

Table 50  
*Mean Change Scores, Standard Deviations, and Confidence Intervals for PYD and CC Subscales as a Function of Transience*

Dependent Variable	Condition	n	$\Delta M$	SD	95% CI	
					LL	UL
Positive Social Development						
Empathic Concern	Low	178	-.02	.78	-.14	.09
	Middle	928	-.18	.87	-.23	-.12
	High	113	-.72	.93	-.89	-.54
Perspective Taking	Low	178	-.02	.82	-.15	.10
	Middle	928	-.08	.86	-.14	-.02
	High	113	-.09	1.02	-.28	.10
Social Rules/Politeness	Low	178	-.06	.66	-.15	.05
	Middle	928	-.09	.68	-.13	-.04
	High	113	-.28	.71	-.42	-.15
Connection to School	Low	178	-.13	.73	-.24	-.02
	Middle	928	-.20	.84	-.26	-.15
	High	113	-.73	.90	-.89	-.56
Connection to Peers	Low	178	-.16	.88	-.29	-.03
	Middle	928	-.17	1.02	-.24	-.11
	High	113	-.41	.96	-.59	-.23
Values Diversity	Low	178	.02	.85	-.10	.15
	Middle	928	-.08	.84	-.14	-.03
	High	113	-.23	.91	-.40	-.06
Personal Values	Low	178	.01	.82	-.11	.13
	Middle	928	-.10	.79	-.15	-.05
	High	113	-.31	.94	-.48	-.13
Classroom Climate						
Positive Behaviour and Classroom Supportiveness	Low	178	-.10	.83	-.23	.02
	Middle	928	-.24	.86	-.30	-.19
	High	113	-.54	.87	-.71	-.38
Negative Behaviour	Low	178	.24	1.27	.06	.43
	Middle	928	.35	1.11	.28	.42
	High	113	.50	1.27	.26	.74
Safety at School	Low	178	.03	.95	-.11	.18
	Middle	928	-.11	1.07	-.17	-.04
	High	113	-.38	.96	-.56	-.20
Enjoyment of School	Low	178	-.29	.97	-.44	-.15
	Middle	928	-.33	1.12	-.30	-.26
	High	113	-.85	1.25	-1.08	-.61

Note. CI = confidence interval, LL = lower limit, UL = upper limit.

When the change in mean scores was compared between children from low and high transient schools, there were statistically significant differences in all seven PYD and in three of the four CC subscales. There were no statistically significant differences in the mean scores for the Negative Behaviour subscale (see Table 51). There were also statistically significant differences between children from middle and low transient schools in the Connection to School, Connection to Peers, Values Diversity, Safety at School, and Enjoyment of School subscales. In each case, children from low transient schools reported smaller decreases in mean subscale scores than children in middle and high transient schools. The size of the effect ranged from trivial to small for both the PYD ( $f^2 = .01 - .04$ ) and CC subscales ( $f^2 = .01$  to  $.02$ ).

Table 51

*One-Way Repeated Measures Analysis of Variance Results and Effect Sizes for Pre-Post PYD and CC Subscale Means as a Function of Transience*

Dependent Variable	<i>F</i> (2, 1216)	<i>p</i>	Effect size differences ( $\eta^2$ )	Effect size differences (Cohen's $f^2$ )
Positive Social Development				
Connection to School	21.73	***	.04	.04
Values Diversity	18.84	***	.03	.03
Empathic Concern	15.74	***	.03	.03
Perspective Taking	8.51	***	.01	.01
Personal Values	5.42	***	.01	.01
Social Rules/Politeness	5.13	**	.01	.01
Connection to Peers	2.99	*	.01	.01
Classroom Climate				
Enjoyment at School	11.31	***	.02	.02
Positive Behaviour and Classroom Supportiveness	9.22	***	.02	.02
Safety at School	2.92	*	.01	.01
Negative Behaviour	1.73	.179	$\leq .01$	$\leq .01$

\*\*\* =  $p < .001$ ; \*\* =  $p < .01$ ; \* =  $p < .05$ .

Overall, these results suggested that the degree of transience within a school environment had a statistically significant adverse impact on the self-reported prosocial attitudes and behaviours for children at both the individual and classroom level. The effect was small but worth noting as children from high transience schools reported larger decreases in feelings of enjoyment and connection to their school over the course of the year. They also reported lower scores for character (Values Diversity), caring (Empathic Concern), and positive behaviour in the classroom (Positive Behaviour and Classroom and School Supportiveness) than children from

low transience schools. This indicated that there may be factors within the high transience schools that were having an adverse impact on children's social attitudes and behaviours. Further, it was suspected that a social health promotion programme, like Kiwi Can, may interact with transience level to influence social attitudes and behaviours.

#### **6.16 Results - The Interactive Effect of Transience and School Category on PYD and CC Subscale Means**

Therefore, the decision was made to explore the relationship between transience level and school condition both within and across time using MANOVAs with two fixed factors and one-way RM-ANOVAs. Unfortunately, the output of the initial set of one-way RM-ANOVAs proved very difficult to interpret due to dramatically different cell sizes in this analysis. In fact, three of the test conditions (*control school-high transience, experienced school-low transience, and new school-low transience*) did not have a single participant. In order to address this issue, the decision was made to remove the control groups from the analysis and to compute the effect of an interaction between transience level and school condition on PYD and CC means using the one-way RM-ANOVA technique. The control group was removed because it clustered very heavily in the low transience group and was completely absent from the high transient condition. By removing the control group, the low transience condition was also effectively removed, therefore all interaction results are presented as a function of two transience conditions – high transience and middle transience, and two school conditions – experienced Kiwi Can and new Kiwi Can.

##### **6.16.1 Baseline means – PYD and CC subscales.**

When the baseline means for PYD and CC subscales data were compared for children as a function of transience and school condition, it was evident that they varied both between and within groups (see Table 52 for the means and standard deviations for PYD and CC subscales).

Table 52

*Means and Standard Deviations for PYD and CC Subscales at Baseline as a Function of School Condition and Level of Transience*

Dependent Variable	School Condition	Transience Level	<i>n</i>	<i>M</i>	<i>SD</i>
Positive Social Development					
Empathic Concern	Experienced	Middle	203	4.82	.91
		High	96	5.30	.71
	New	Middle	447	5.12	.89
Perspective Taking	Experienced	High	17	4.31	.96
		Middle	203	4.43	.93
	New	High	96	4.64	.80
Social Rules/Politeness	Experienced	Middle	447	4.68	.86
		High	17	4.26	1.18
	New	Middle	203	5.10	.79
Connection to School	Experienced	High	96	5.40	.58
		Middle	447	5.37	.65
	New	High	17	4.76	1.01
Connection to Peers	Experienced	Middle	203	4.83	.80
		High	96	5.25	.67
	New	Middle	447	5.05	.81
Values Diversity	Experienced	High	17	4.51	.79
		Middle	203	4.70	1.01
	New	High	96	4.71	.83
Personal Values	Experienced	Middle	447	4.82	.99
		High	17	4.05	1.26
	New	Middle	203	4.61	.85
Classroom Climate	Experienced	High	96	5.16	.60
		Middle	447	4.85	.85
	New	High	17	4.10	1.05
Positive Behaviour and Classroom Supportiveness	Experienced	Middle	203	4.76	.84
		High	96	5.14	.76
	New	Middle	447	5.09	.76
Negative Behaviour	Experienced	High	17	4.30	1.23
		Middle	203	4.50	.84
	New	High	96	4.88	.72
Safety at School	Experienced	Middle	447	4.77	.81
		High	17	4.25	.98
	New	Middle	203	3.96	1.02
Enjoyment of School	Experienced	High	96	3.80	1.15
		Middle	447	3.63	1.11
	New	High	17	3.84	.95
	Experienced	Middle	203	4.66	.93
		High	96	5.13	.83
	New	Middle	447	4.92	.93
	Experienced	High	17	3.87	.86
		Middle	203	4.64	1.19
	New	High	96	5.19	1.10
		Middle	447	5.05	.98
		High	17	4.04	.75

Small statistically significant differences were identified at baseline for the PYD subscales, Pillai's  $V = .04$ ,  $F_{(7, 1207)} = 7.10$ ,  $p < .001$ , as a function of the interaction between transience level and school condition (see Table 53 for the Between-Subjects

Effects output from the MANOVA). The size of the effect was trivial for Perspective Taking and Connection to Peers ( $f^2 < .02$ ) and very small ( $f^2 = .02$  to  $.03$ ) for Empathic Concern, Social Rules/Politeness, Connection to School, Values Diversity, and Personal Values. For the CC baseline data, statistically significant differences also arose for three subscales – Positive Behaviour and Classroom Supportiveness, Safety at School, and Enjoyment of School, as a function of the interaction between school condition and transience,  $V = .04$ ,  $F_{(4, 1210)} = 12.02$ ,  $p < .001$ . The effect sizes for these differences were trivial for Positive Behaviour ( $f^2 < .02$ ) and very small for Safety at School and Enjoyment of School ( $f^2 = .02$  to  $.03$ ). In each case, new Kiwi Can schools under the high transience condition reported significantly lower mean scores for the subscales than did experienced children under the high transience condition.

Table 53

*MANOVA output for Effect of Interaction between School Condition and Transience Level on Baseline PYD Mean Scores*

Dependent Variable	$F(1, 1213)$	$p$	Effect size differences $\eta^2$	Effect size differences Cohen's $f^2$
Positive Youth Development				
Personal Values	31.87	***	.03	.03
Empathic Concern	30.44	***	.02	.03
Values Diversity	28.87	***	.02	.02
Social Rules/Politeness	24.29	**	.02	.02
Connection to School	20.07	***	.02	.02
Connection to Peers	9.26	**	.01	.01
Perspective Taking	8.99	**	.01	.01
Classroom Climate				
Safety at School	37.94	***	.03	.03
Enjoyment at School	28.71	***	.02	.02
Positive Behaviour and Classroom Supportiveness	16.33	**	.01	.01
Negative Behaviour	1.55	.214	$\leq .01$	$\leq .01$

\*\*\* =  $p < .001$ ; \*\* =  $p < .01$ ; \* =  $p < .05$ .

### 6.16.2 Change in subscale means - PYD and CC.

To calculate the effect of transience and school condition on the change in PYD and CC means over time, four dichotomous between-subjects variables were created. They were – new Kiwi Can school-middle transience (*new-middle*), new Kiwi Can school-high transience (*new-high*), experienced Kiwi Can school-middle transience

(*experienced-middle*), and experienced Kiwi Can school-high transience (*experienced-high*).

The results of the one-way RM-ANOVAs showed that the interaction had a statistically significant but small effect ( $f^2 = .02$  to  $.09$ ) on mean change in six PYD subscales. The effect was largest for Connection to School,  $V = .08$ ,  $F_{(3, 759)} = 21.94$ ,  $p \leq .001$ , followed by Values Diversity,  $V = .04$ ,  $F_{(3, 759)} = 10.22$ ,  $p \leq .001$ , Empathic Concern,  $V = .03$ ,  $F_{(3, 759)} = 7.97$ ,  $p \leq .001$ , Social Rules/Politeness,  $V = .03$ ,  $F_{(3, 759)} = 7.13$ ,  $p \leq .001$ , Connection to Peers,  $V = .02$ ,  $F_{(3, 759)} = 6.29$ ,  $p = .001$ , and Perspective Taking,  $V = .02$ ,  $F_{(3, 759)} = 5.59$ ,  $p = .001$ . The interaction between school condition and transience level had a statistically significant effect on Personal Values,  $V = .01$ ,  $F_{(3, 759)} = 3.32$ ,  $p \leq .001$ , however, it was trivial in size. For the classroom climate subscales the interaction between school condition and transience level also had a significant but small ( $f^2 = .02$  to  $.06$ ) effect on the change in mean scores for Positive Behaviour and Classroom Supportiveness,  $V = .06$ ,  $F_{(3, 759)} = 16.56$ ,  $p \leq .001$ , Enjoyment of School,  $V = .03$ ,  $F_{(3, 759)} = 7.99$ ,  $p \leq .001$ , and Negative Behaviour,  $V = .02$ ,  $F_{(3, 759)} = 4.74$ ,  $p \leq .001$ . There were no statistically significant differences in mean change for the Safety in School subscale. According to the pairwise comparison output for each statistically significant interaction, the means for all seven PYD and three CC subscales decreased more for the new-high transience schools than for the new-middle, new-high, and experienced-high transience schools.

See Table 54 for the change in mean values for each subscale and Table 55 for the results of the one-way RM-ANOVAs.

Table 54

*Mean Change Scores, Standard Deviations, and Confidence Intervals for PYD and CC Subscales as a Function of Transience Level x School Condition*

Dependent Variable	Condition	n	$\Delta M$	SD	95% CI	
					LL	UL
Empathic Concern	Middle-Experienced	197	-.21	.93	-.34	-.08
	Middle-New	373	-.15	.86	-.23	-.07
	High-Experienced	96	-.64	.91	-.82	-.45
	High-New	97	-1.18	.96	-1.67	-.69
Perspective Taking	Middle-Experienced	197	-.03	.93	-.16	.10
	Middle-New	373	-.08	.83	-.15	.00
	High-Experienced	96	-.06	1.04	-.27	.15
	High-New	97	-.29	.94	-.77	.19
Social Rules/Politeness	Middle-Experienced	197	-.14	.75	-.25	-.04
	Middle-New	373	-.04	.61	-.10	.00
	High-Experienced	96	-.21	.71	-.36	-.07
	High-New	97	-.69	.55	-.98	-.41
Connection to School	Middle-Experienced	197	-.22	.88	-.35	-.10
	Middle-New	373	-.10	.77	-.17	-.03
	High-Experienced	96	-.61	.85	-.79	-.44
	High-New	97	-1.36	.93	-1.84	-.88
Connection to Peers	Middle-Experienced	197	-.19	1.05	-.33	-.01
	Middle-New	373	-.11	.98	-.20	-.02
	High-Experienced	96	-.28	.85	-.45	-.11
	High-New	97	-1.13	1.23	-1.76	-.49
Values Diversity	Middle-Experienced	197	-.15	1.01	-.30	-.01
	Middle-New	373	-.05	.79	-.13	.02
	High-Experienced	96	-.16	.81	-.12	.00
	High-New	97	-.66	1.28	-1.32	.00
Personal Values	Middle-Experienced	197	-.12	.86	-.24	.00
	Middle-New	373	-.10	.74	-.17	-.03
	High-Experienced	96	-.25	.98	-.45	-.05
	High-New	97	-.65	.60	-.96	-.35
Positive Behaviour and Classroom Supportiveness	Middle-Experienced	197	-.22	.88	-.34	-.09
	Middle-New	373	-.17	.84	-.25	-.10
	High-Experienced	96	-.36	.74	-.51	-.20
	High-New	97	-1.61	.80	-2.02	-1.20
Negative Behaviour	Middle-Experienced	197	.10	1.19	-.07	.27
	Middle-New	373	.38	1.10	.28	.48
	High-Experienced	96	.43	1.30	.16	.69
	High-New	97	.92	1.00	.41	1.44
Safety at School	Middle-Experienced	197	-.06	1.20	-.23	.10
	Middle-New	373	.00	1.01	-.10	.00
	High-Experienced	96	-.31	.98	-.51	-.11
	High-New	97	-.77	.74	-1.15	-.39
Enjoyment of School	Middle-Experienced	197	-.36	1.36	-.55	-.17
	Middle-New	373	-.27	1.04	-.36	-.17
	High-Experienced	96	-.79	1.29	-1.05	-.53
	High-New	97	-1.18	.93	-1.65	-.70

Table 55

*One-way RM-ANOVA output for Effect of the Interaction between School Condition and Transience Level on Change in PYD and CC Mean Scores*

Dependent Variable	<i>F</i> (3, 759)	<i>p</i>	Effect size differences $\eta^2$	Effect size differences Cohen's $f^2$
Positive Social Development				
Connection to School	21.94	***	.08	.09
Values Diversity	10.22	***	.04	.04
Empathic Concern	7.97	***	.03	.03
Social Rules/Politeness	7.13	***	.03	.03
Connection to Peers	6.29	***	.02	.02
Perspective Taking	5.59	**	.02	.02
Personal Values	3.32	**	.01	.01
Classroom Climate				
Positive Behaviour and Classroom Supportiveness	16.56	***	.06	.06
Enjoyment at School	7.99	***	.03	.03
Negative Behaviour	4.74	**	.02	.02
Safety at School	2.44	.064	.01	.01

\*\*\*=  $p < .001$ ; \*\*=  $p < .01$ ; \* =  $p < .05$ .

### **6.16.3 Summary of change in PYD and CC subscale means as a function of transience level and school condition.**

There was a statistically significant effect of school condition and transience level on self-reported (social) Competence, Connection to School, Connection to Peers, Caring and Compassion, and Character over time. There was also a statistically significant effect on the change in self-reported incidence of positive, prosocial attitudes and behaviour within the classroom, and reported enjoyment while at school. In each case, children from highly transient schools who were new to the Kiwi Can programme reported significantly larger decreases in mean scores for the same ten subscales across time than children from highly transient schools who had participated in the Kiwi Can programme for two years or more. Further, children from highly transient-new schools also reported significantly larger decreases in mean scores for all ten subscales than children from middle transient schools. These results suggest two things – firstly, that the degree of school-level transience had an effect on the incident of self-reported prosocial attitudes and behaviours amongst primary and intermediate students from low-decile schools in this study, and secondly, that sustained participation in a positive youth development programme, like Kiwi Can, may attenuate some of the negative effects of a high transience school environment on children's social attitudes and behaviours.

## Chapter 7: Discussion

### 7.1 Overview of Results

The research in this thesis explored a number of specific, inter-related aspects of health education curriculum design and delivery in the New Zealand schooling context. Specifically, it investigated teacher's perceptions of the delivery of the health education curriculum in primary schools (Study 1), and it used a multi-method design to evaluate the implementation and impact of participation in an externally provided, social health promotion programme, Kiwi Can, on the social attitudes and behaviours of students in low-decile primary and intermediate schools (Study 2 and Study 3). The Kiwi Can programme evaluation was conducted over a period of two years (2012 to 2013).

At the end of the two year data collection and analysis process, the results from this PhD study proved both informative and intriguing. The results from Study 1 suggested that the current social health education curriculum was being delivered only infrequently, if at all, and that the main barriers to delivery were a lack of time and a lack of resources. The results from Study 2 indicated that the Kiwi Can programme was being implemented to a high standard, with fidelity, and adherence to participant responsiveness in all participating Kiwi Can regions ( $n = 8$ ). The results from Study 2 also suggested that there were no large, statistically significant differences in programme implementation as a function of year level, sex, ethnicity, or school conditions across the regions. The results from Study 3, the outcome evaluation, were varied and must be considered in four parts – baseline PYD and CC results, pre-post PYD and CC results, transience and PYD/CC results, and transience and school condition results.

At baseline, the mean scores for all seven PYD subscales ( $M > 4.00$ ) and all four CC subscales ( $M > 3.55$ ) were high. Further, there were no large, statistically significant differences in PYD or CC subscale means as a function of school condition, year level, sex, or ethnicity. Also, the results of the pre-post outcome evaluation surveys indicated that all PYD and CC subscale mean scores decreased for children in Kiwi Can and control schools from Time 1 to Time 2. When the results were explored as function of school condition, children engaged in the Kiwi Can programme

experienced a small but statistically significant decrease in the mean scores for all individual PYD and all positively-oriented CC factors over time. The decrease in mean scores across four subscales – Empathic Concern, Connection to School, Values Diversity, and Safety in School, was significantly larger for children from experienced Kiwi Can programme schools than for those from new Kiwi Can and control schools. Given that the Kiwi Can programme was designed to promote positive social attitudes and behaviours, these results were not expected. They led to the design of an exploratory study that investigated the impact that school level transience had on PYD and CC subscale mean scores at baseline and over time.

The results of the exploratory study indicated that the self-reported PYD and CC subscale mean scores were influenced by the degree of transience in the schooling environment. At baseline, mean scores for individual (PYD) and classroom-level factors (CC) were lowest in schools with the highest rates of transience ( $\geq 32.89\%$ ). Across time, mean scores decreased for all three transience conditions (low, middle, and high), however, they decreased more substantially in the high transience condition than the low transience condition. This suggested that a highly changeable, or transient, school environment contributed to lower prosocial attitude and behaviour scores in non-transient students over time. It also suggested that there may be characteristics, or factors, in a highly transient schooling environment that interact with children to create negative social attitudes and behaviours over time. The study did not, however, formally identify these environmental factors, nor did it explore the interplay between the individual and the environmental factors that led to these results. Further empirical studies are therefore required before conclusions can be drawn about the cause of the poor social attitude and behaviour scores reported by non-transient children in highly transient school settings. In this PhD, a final empirical study was conducted to explore the interaction between school level transience and programme condition on children's social attitudes and behaviours.

The results of the final interaction study indicated that the degree of transience in the school interacted with the degree of experience a school had with the Kiwi Can programme. In the middle transient school condition, the decrease in mean scores was significantly larger for students in experienced Kiwi Can than in new Kiwi Can schools. Under the highly transient condition, however, the decrease in mean PYD

and CC scores was significantly larger for children in the new Kiwi Can schools rather than the experienced schools. This result indicated that sustained exposure to the Kiwi Can programme (i.e., > 2 years) may have attenuated some of the negative effects that a highly transient schooling environment had on the social attitudes and behaviours of vulnerable, non-transient students in low-decile schools. The results of this final study highlighted the fact that participation in a social health promotion programme, like Kiwi Can, must be sustained over a lengthy period of time in order to realise a social health benefit. In this study, a single year or two of participation was not long enough for students to experience a positive change in their social attitudes or behaviours. This finding is consistent with many models of behaviour change, most notably with Ajzen's theory of planned behaviour (TPB; Ajzen, 1991, 2011). According to the continuum-based TPB model, changes in attitudes and behaviours, particularly those involving deeply embedded beliefs, take time because of the inherently complex nature of behaviour change. Ajzen (1991) suggested that behaviours will not change unless individuals develop both the intention and the perceived ability (behavioural control) to successfully perform the action. The intention is a product of an individual's attitude towards the behaviour and the subjective norms associated with it. Subjective norms are a function of the expectations that significant others have about the behaviour, and are weighted by a motivation to conform (Mullen, Hersey, & Iverson, 1987). In the context of social behaviour change, this theory suggests that children will engage in prosocial behaviour only after a lengthy period of time during which they evaluate the importance attributed to the behaviour by others in their social world, assess the consequences of performing that action, develop positive attitudes about it, and acquire the resources and skills needed to perform the pro-social behaviour. Ajzen (1991) argues that this can take years.

Given the inherent complexities and length of time needed to change children's behaviour, social health promotion programmes, like Kiwi Can, might need to be adopted and enculturated into low-decile schooling environments for years before the positive effects of participation can be fully realised. The results also suggest that the degree of transience within the school can affect children's attitudes and behaviours, and that they must be carefully considered when designing, implementing, and evaluating the efficacy of a school-based health promotion initiative. This raises

questions about the current practice of educators, programme design and delivery teams, and programme evaluators. From a theoretical perspective, these results also raise important questions about current understandings of positive developmental theory and its application in highly transient, low-decile primary and intermediate school settings.

## **7.2 Implications**

### **7.2.1 Implications for primary school health educators and external providers of health education programming in New Zealand.**

From a pragmatic perspective, the results from this thesis have a number of important implications for practice in programme design, delivery, and evaluation in school settings. The most notable of these is the apparent impact that transience has on individual social attitudes and behaviours, as well as on the climate of the classroom for non-transient children in low-decile schools. It is clear that non-transient children in highly transient environments struggle to forge healthy social relationships, feel connected to their peers or schools, behave in prosocial ways, or show caring and compassion for others. They also tend to feel less safe in their schools and enjoy their schooling experience less than those in more stable, less transient environments. While these results have been reported by researchers exploring the impact of transience on the transient student, they had yet to be explored amongst the non-transient population.

This PhD study is an exception. The entire sample of children in Study 3 was, in effect, immobile. They had all attended their school for the full academic year and provided the researcher with results at Time 1 and Time 2. Given what is known about the disruptive nature of student mobility, it is not surprising that these children struggled to forge and maintain meaningful relationships with peers who were routinely moving in and out of their lives. While it is likely that staff at these highly transient schools worked with children to establish a stable culture and mitigate against the constant disruption, the findings from this PhD suggest that their efforts were not sufficient. The results also suggest that targeted programming helped overcome the negative effects of transience provided sufficient time was permitted for the additional assets to be developed among the immobile student community. For

schools, and programme providers like Kiwi Can, this points out significant new directions for the current approach. It also raises awareness that there is a need for strategies that can better address the negative effects of high mobility on children's social health.

The Education Review Office in New Zealand [ERO] has undertaken preliminary work in this area. In 2007, representatives of ERO conducted a small study of effective practices for addressing transience in primary schools. In their summary report (ERO, 2007), the authors listed the following examples of good practice –

- (i) promote the notion of the school as one family, with emphasis on aroha, tautoko and whanaungatanga
- (ii) provide orientation programmes, buddy systems, and peer support groups to help new students settle into the new school
- (iii) supply uniforms for transient students
- (iv) support students' families with links to local support services
- (v) build connection with students' families

While the strategies have proven to be efficacious for transient students, they do not address the needs of immobile students. They are also quite general and fail to take into account the diversity of individual experiences. Given the results of the Kiwi Can outcome evaluation, this is both a necessary and important goal. In the Kiwi Can outcome evaluation, children with different demographic characteristics responded differently to aspects of the social health questionnaire. In fact two groups in particular reported markedly lower mean scores for individual social development (PYD) and classroom climate (CC) than their peers. These groups were boys, and Year 8 students. While disappointing, these results were not unexpected given that Archer (2009), Toldos (2005), and von Collani and Werner (2005) have previously reported a significantly higher incidence of antisocial behaviour in male children than female children. In these studies, the authors suggested that males reported a higher incidence of antisocial behaviour because of increased exposure to, and familiarity with, the behaviour. It follows then, that boys in this study may have reported lower positive social attitude and behaviour subscale mean scores because they were the perpetrators and/or victims of antisocial behaviour on a more regular basis than their female counterparts. This may have affected their feelings of competence,

connection, caring/compassion, and their perception of the classroom environment as a positive social space. It is also possible that the males in this study reported lower PYD and CC subscale scores than females because they were simply less engaged with the Kiwi Can programme, and therefore less likely to internalise the messages and change their attitudes and behaviours. Further, the lower subscale mean scores may be a by-product of the way in which children are socialised and expected to embody and actualise gender in New Zealand. Typically, boys are expected to display their malehood through strong and often aggressive acts of masculinity, while girls are generally expected to be gentle, empathic, caring, and sensitive – characteristics that were being measured in this study. Therefore, the lower scores reported by boys for caring, compassion, connection, social competence and character may indicate that the boys had either refrained from engaging in typically female behaviour, and/or underreported their involvement in prosocial behaviour because they had been socialised to believe that this feminine behaviour was not socially acceptable or appropriate to demonstrate.

From the perspective of age and year level, the lower scores for Year 8 children reported in Study 3 are likely to be a by-product of the impending transition from the safety of the middle school environment to the unknown life of a secondary school student. As they anticipated leaving their peers and school behind, they may have consciously (or unconsciously) distanced themselves, thus resulting in lower levels of commitment to maintaining and establishing positive social relationships with their peers and the school.

The results from this thesis have implications for the design of school-based social health education and social health promotion programmes, like Kiwi Can. In fact, they suggest that it may be advantageous to re-design social health education programming to focus specifically on the development of social assets (connection, competence, caring and character) that enable non-transient children in highly transient settings to cope with the unstable nature of their social environment. Through the programme, participants should be taught how to build and maintain strong social networks and access the social supports they require both within the classroom and the wider school environment. Participants must also be taught to develop positive, prosocial attitudes and to engage in prosocial behaviours despite the

highly changeable and transient nature of their environment. From a design perspective, the results suggest that the social health programme curriculum in highly transient schools must include lessons that equip students with the skills required to maintain healthy social relationships during periods of transition and to challenge current gender stereotypes. These social health promotion programmes could be offered by either external agencies, such as FYD, or by health education teachers who are trained in the university setting to deliver a comprehensive social health promotion programme. This would involve a significant commitment on the part of universities and teacher training providers in New Zealand.

### **7.2.2 Implications for curriculum and teacher professional development in health education.**

At present, very little time and resourcing is given to educating the incoming HPE workforce about the health education curriculum. According to Smith and Philpot (2011), students receive as little as 18 hours of formal training in health education prior to commencing their professional careers as teachers. The results from this PhD study indicate that many of these teachers then enter the workforce and shy away from delivery in areas that are unfamiliar and/or disinteresting to them – areas such as mental, emotional, and social health education. These teachers also shy away from delivery in health education when they feel that they lack the time or the resources necessary to deliver it. These results are not unique to the New Zealand context.

Such findings highlight the negative impact that time and resourcing can have on the delivery of school-based health education. They also raise awareness of the effect that these perceived barriers have on practice, and the reality that health education teachers in New Zealand and overseas are often forced to make instructional choices. According to the results of Study 1, health education teachers in New Zealand choose to respond to these barriers by selectively favouring instruction in one area of the health education curriculum – physical health. This action raises philosophical questions about how health education practitioners value and prioritise learning in the different health dimensions, as well as questions about how to alleviate the negative effect of time and resourcing on health education teacher practice. One possible mechanism explored in this PhD was the introduction of a fully resourced social health promotion programme, like Kiwi Can, offered by an external provider. The

Kiwi Can programme came with a complete set of differentiated lesson plans that were designed for use with children from five to 14 years of age, and that were, most importantly, underpinned by a positive developmental approach to social health education. The results of this study provide support for continued exploration of alternative methods for social health promotion, particularly in low-decile primary and intermediate schools.

In the absence of adequate resourcing, New Zealand health education teachers and New Zealand children continue to miss out on unique and meaningful opportunities to engage with the health education curriculum, develop healthier attitudes, and adopt healthier behaviours. This may adversely affect the social health and wellbeing of our most vulnerable population – a population that may already struggle with bullying and violent behaviour at home and school, and is disadvantaged by not having extra assets to deal with a swirling and unstable social context.

### **7.2.3 Implications for positive youth development theory.**

From a theoretical perspective, the findings from this thesis also raise a number of important questions about the integrity of positive youth development theory and the application of Lerner's 5C's theory in the primary school context in New Zealand. Firstly, the results of the pre-post survey data suggest that participation in a social health promotion programme is associated with poorer, rather than greater, expression of the 'C's. In the Kiwi Can outcome evaluation, experienced schools reported lower mean scores for caring/compassion, connection, social competence, and character than new and/or control schools at Time 2. These results were replicated for the classroom climate subscales. Children from experienced schools reported exposure to more antisocial behaviours and attitudes within the school classroom than those in new or control schools. These findings are in contrast to those of Jellic, Bobek, Phelps, Lerner, and Lerner (2007) and Phelps, Balsano, Fay, Peltz, Zimmerman, and Lerner (2007), who found that participation in positive youth development programmes was associated with reduced risk taking behaviour. It is possible that the results from this study differed from those previously reported, as this was the first time that Lerner's framework had been applied in a New Zealand setting. Perhaps there are significant cultural differences at play that are influencing the way in which children from New Zealand experience positive youth development programming and

respond. It is also possible that the results reflect the fact that previous work has typically been undertaken with adolescents, while the population in this study was distinctly younger (Years 6-8). It must be noted however, that this not a strong possibility as both Jelcic et al. and Phelps et al. conducted work with children as young as 10 years of age and found evidence in support of Lerner's typology. The population in these American studies however, was not drawn exclusively from low-decile neighbourhoods. In fact, only one-third of the participants in the 4-H studies came from homes with low average incomes ( $\leq$  \$34, 999 USD) (Lerner et al., 2005). This highlights the need for additional work to be done in both New Zealand and abroad in order to further explore the impact that socioeconomic status has on the development and expression of prosocial behaviour in children.

The findings from this PhD thesis also question Benson's assertion that the school environment could serve as an asset or contributing factor in the positive development of youth. According to Benson's 40 developmental assets theory, a caring school climate serves as one of the external assets that can facilitate healthy personal growth and development (Benson et al., 1998). Researchers such as Eccles and Roeser (1999), Kuperminc, Leadbeater, Emmons, and Blatt (1997) and Roeser, Eccles, and Sameroff (2000) have also reported that the classroom climate of primary and intermediate schools can strongly influence student adjustment and behaviour across multiple domains including the social dimension. It was, therefore, expected that a positive classroom climate might serve as an asset, and possibly predict scores for individual social attitudes and behaviours. The results from this PhD study did not provide explicit support for a relationship (interactive or cross-lagged effect) between individual social development and the social climate of a school classroom. However, given that this study was measuring positive youth development and classroom climate at the individual level, and school climate is only one of 40 assets in the framework, it is not possible to draw any significant conclusions about the merit and applicability of Benson's model in low-decile schools in New Zealand.

### **7.3 Successes and Challenges**

Overall, this PhD study was able to accomplish all of its stated aims – to investigate the delivery of the health education curriculum in New Zealand primary schools, with particular focus on the social health dimension, and to evaluate the implementation

and outcome of participation in an externally provided, social health education and promotion programme on the social attitudes and behaviours of vulnerable children in low-decile schools. It was also able to address each of the stated research questions, and to explore the relationships between degree of transience and social health outcomes for a population that is grossly understudied and underserved – non-transient school children. In fact, this thesis draws our attention to the unique and significant impact that a highly transient classroom and schooling environment is having on the children who are exposed to it. It represents a first attempt to quantify the impact that transience is having on the social health outcomes for this population and suggests that participation in a positive social health promotion programme may be one way of attenuating the effect.

It must be noted however, that the design of the current PhD study was not without challenge or limitation. In fact, four notable challenges have been identified that may limit the generalisability of the results from one of the three studies, the outcome evaluation of Kiwi Can. Firstly, the outcome evaluation data were collected at two discrete time points only. In the absence of a third (or fourth) data collection phase, it was not possible to generate a latent growth curve model that was more sensitive to changes in individual responses over time. Secondly, the sample sizes were too small in some cases to support the design of a hierarchical linear model, and therefore it was not possible to directly account for the nested structure of the data. Small sample sizes also became an issue when the interaction between level of transience and programme condition was explored. In these analyses, two of the cells under investigation (control-high transience, and experienced-low transience) were completely empty. This generated covariance matrices that were not homogeneous. Thirdly, it must be noted that the data for the outcome evaluation of Kiwi Can was collected by school teachers and programme leaders rather than the researcher. This may have adversely affected the quality of the data that were collected both pre-intervention and post-intervention. Finally, the data were collected over a single year of schooling. One year may not have been a sufficient period of time to establish the true effect of resource accumulation on children's prosocial attitudes and behaviours.

Despite the challenges associated with the nature of the data collection and analysis for the outcome evaluation, the researcher is confident that the data collected were

valid and that the results generated were a reliable and accurate account of the lived experiences and social health status of primary and intermediate school aged children from low-decile communities throughout New Zealand. In order to provide additional support for this assertion however, further study is needed.

#### **7.4 Future Directions**

Prior to embarking on future work in this area, it is important to critically reflect on the methodology employed in the outcome evaluation of Kiwi Can and the results achieved. This is particularly important given that the results were not consistent with research results from studies undertaken in the United States under similar conditions. After much consideration and thoughtful reflection, the researcher believes that the results of the outcome evaluation of Kiwi Can may have been influenced by a design flaw, participant bias, and/or 'other' confounding factors. In fact, it is possible that the lower scores for PYD and CC factors at Time 2 (end of the year) were a product of decreased student motivation and fatigue. According to Eccles, Wigfield, Midgeley, Reuman, MacIver, and Feldlaufer (1993) and Smith (2004), primary, intermediate, and secondary school students can experience a significant decrease in energy and motivation at the end of an academic school year. They can also experience a decrease in mastery, academic performance, and self-motivation, as they age (Midgeley & Edelin, 1998; Murphy & Alexander, 2000; Yeung & McInerney, 2005). Also, it is not uncommon for primary school-aged children to become more realistic about their experiences and to judge them with a more objective lens as they age (Yeung, Lau, & Nie, 2011). For some, this might have involved a transition from being poised and excited about the beginning of a new school year to feeling disappointed by what failed to transpire over the preceding 10 months. This perceived failure may have altered their mood or attitude when completing the social health questionnaire at Time 2. On the basis of these findings, it is possible that the lower individual, prosocial and classroom climate scores may have been the result of child fatigue, apathy, or loss of objectivity, at the end of the 2013 school year.

While it is probable that the survey results were influenced by participant characteristics, it is also possible that they were influenced by the design of the measurement tool (the social health questionnaire). In offering students a positively-packed item-response format, it is possible that the researcher inadvertently conflated

the baseline results by restricting the range of responses. Given that the average score at baseline was a '4' (out of 6), it is plausible that children responded so positively at Time 1 that they had no opportunity to indicate significant positive change at the end of the year. This restriction of range issue has been discussed at length in the literature by Dunham and Davison (1991), Frisbie and Brandenburg (1979), and Lam and Klockars (1982). It has also been described by adolescent health researchers in the United States, who reported that mean scores for behaviour tend to be positively biased amongst youth who agree to participate in longitudinal research studies (Baltes, Reese, & Nesselroade, 1977; Schaie & Strother, 1968).

The results achieved in the Kiwi Can outcome evaluation may have also been a by-product of the manner in which classroom climate was measured in the course of this research. While many researchers use instruments for measuring classroom climate that focus on aspects of teacher support, student autonomy, rule clarity, and rule fairness within the school environment, the instrument adopted for use in this thesis focussed exclusively on the social attitudes and behaviours of children, though it was intentionally designed for this purpose as the intent was to measure the social attitudes and behaviours of children within the classroom setting. It was however quite limited in scope, and possibly failed to capture elements of the classroom climate that were most instrumental for predicting children's adjustment and behaviour scores (i.e., teacher support, teacher behaviour, and school rules). In order to test this theory, the outcome evaluation of Kiwi Can would have to be repeated using a more traditional measure of classroom climate such as Trickett and Moos' (1973) Classroom Environment Scale, or the tool adopted for use in this PhD study, The School Climate Survey, in its entirety (Ding et al., 2011).

Given what has been learned through this process it could be suggested that further studies explore the impact that participation in a school-based, social health promotion programme has on children's attitudes and behaviours using a minimum of three data collection points and a larger number of participating schools. These increases would enable the generation of more complex models of change over time (Latent Growth Models), using SEM approaches and employing hierarchical linear modelling (HLM) techniques that would be able to account for the nested structure of the data. To further strengthen future study designs, the number of year levels tested

could be expanded from three (Years 6-8) to five (Years 4-8). This would facilitate a more comprehensive analysis of the differences in positive social development and classroom climate outcomes by year level.

While a study of this nature would provide data on the general impact of participation in a social health programme, it would not specifically investigate the impact of a health promotion programme on the attitudes and behaviours of children under varying conditions of transience. Given that the results from this PhD suggest that there may be an interactive effect of transience and school condition on the social health outcomes for immobile students in highly transient schools, it is important that this finding be explored further. Ideally, it would be explored among primary and intermediate students from Kiwi Can schools in low-decile communities.

A possible study design could involve collecting cross-sectional data from Kiwi Can programme leaders, classroom teachers, children and their families, at four low-decile, high transience primary schools in South Auckland. The schools would be purposively selected on the basis of their active participation in Kiwi Can and level of transience (>30% transience over the previous school year). Data for the study could be collected in two parts. In Part one, separate semi-structured focus group interviews could be conducted with classroom teachers, Kiwi Can leaders, children from Year 4 to Year 6 classes, and families from each of the four participating schools. The data from the first part, in conjunction with an inspection of relevant research literature, could then be used to develop a revised 12 week Kiwi Can programme that would provide strategies to mitigate the impact of transience on classroom climate and social health. The curriculum could be written in collaboration with schools and the Kiwi Can team. In Part two, the enhanced programme would be trialled with all Year 4-6 classes from the four primary schools. The fidelity of implementation and programme impact would then be evaluated. Using a pre-post test design, all Year 4-6 children who participated in the programme would be invited to complete previously developed self-report surveys assessing class climate and social health. These results would be compared with data gathered from highly transient schools surveyed in 2013 during this PhD study. Fidelity would also be assessed through a second round of focus groups with Kiwi Can leaders, teachers, children, and families using the same sampling frame outlined previously. In this phase, participants would

be asked to comment on the enhanced programme and its impact. The overarching goals of a project like this would be to explore the risk and protective factors that influence the social health of primary school children from high transience, low-decile schools, and to design and evaluate the efficacy of a modified version of Kiwi Can specifically developed to mitigate the impact of transience on social health. It is hoped that a project of this nature would provide robust evidence of the efficacy of a targeted social health promotion programme, and would address some of the challenges and logistical limitations faced by the researcher in this PhD study.

## **7.5 Contributions**

The findings of the three sequential studies in this PhD are significant and substantive, and contribute to understandings of health education curriculum design and delivery in a number of valuable and unique ways. They also contribute to the scholarship in the areas of programme evaluation methodology and social health promotion for non-transient children.

From a programme design and delivery perspective, the results of this thesis provide robust evidence of the manner in which the primary school health education curriculum is being delivered in New Zealand schools. This work is substantive and unique because it is only the second study of its kind in New Zealand to explore the implementation of the health education curriculum, and the first to focus specifically on teacher beliefs about the importance and benefit of delivery in the social health dimension. It is also the first to consider the barriers to, and use of, external providers of social health promotion programming in the delivery of the curriculum. The results of this study are useful for health education curriculum writers and practitioners who are interested in understanding how the curriculum is perceived, and how it is delivered. It also highlights the fact that teachers are failing to deliver a comprehensive health education programme and signal a need for a shift in thinking and resourcing for health education in New Zealand schools. The results from the Kiwi Can programme evaluation studies also contribute to our understandings about health education delivery in school settings and provide evidence of the positive impact that an externally provided programme can have on child social health and wellbeing. Kiwi Can is the first externally provided social health promotion

programme to be empirically evaluated in low-decile primary and intermediate schools in New Zealand.

From an evaluation methodology perspective, the Kiwi Can implementation and outcome evaluation studies contribute to the field as they are first to use confirmatory factor analysis and structural equation modelling techniques to evaluate positive youth development programme delivery and impact in the New Zealand context.

From a social health perspective, the results of this PhD are unique and substantive. They present evidence that transience can have a negative effect on the social attitudes and behaviours of vulnerable children in low-decile school settings who are not themselves transient. Attention is drawn to the plight of the immobile student – a topic rarely covered in the literature, and never discussed in New Zealand. This is the first study of its kind to identify a relationship between transience and social wellbeing among non-transient children. The findings raise important questions about how to address the negative and disruptive effects of transience, and they highlight the need for additional work to be undertaken to further understand and elucidate the health impacts of membership in a highly transient school community. This PhD also highlights the different lived experiences of children in social health promotion programmes, and points to the need for differentiating programme delivery, and monitoring impact.

Lastly, and most importantly, this PhD represents an important first step in articulating a vision for social health promotion amongst vulnerable children in New Zealand. As has been expressed by the United Nations, the New Zealand Government, and the New Zealand public, promoting child health, particularly amongst those from the most vulnerable and challenging of circumstances, is an important and necessary goal, for...

He who has health has hope, and he who has hope has everything.

## Appendix A

### Memorandum of Understanding

Between the Foundation for Youth Development and the University of Auckland

#### RE: KIWI CAN PROGRAMME OUTCOME EVALUATION

The Foundation for Youth Development (FYD) operates a school-based values and life skills called *Kiwi Can*. Given that the FYD wishes for the *Kiwi Can* program to be evaluated so as to inform its improvement and given that Ms. Rachel Williams wishes to complete a research doctorate in positive youth development at The University of Auckland, it is understood that there is an opportunity to collaborate for mutual benefit. The purpose of this memorandum is to set out mutually agreed guidelines for how the evaluation of *Kiwi Can* will be implemented for benefit to both parties.

The University of Auckland party to this agreement consists of Ms. Williams, a PhD candidate and Doctoral Scholar at the University of Auckland, Faculty of Education and her two supervisors (Associate Professor Gavin Brown and Dr. Pat Bullen. Both supervisors are permanently appointed research academics at the University of Auckland with expertise in evaluation methodology and positive psychology. For the purposes of this document, they will be referred to collectively as the ‘academic research team.’

For the purposes of the research evaluation, the FYD party will be represented by: the Manager of Programme Development, Ms. Victoria Hood; the *Kiwi Can* Programme Developer, Ms. Brownyn Blair; and the Manager of Research and Evaluation, Ms. Julie Moore. The outcome evaluation of *Kiwi Can* will be conducted in accordance with the following guidelines:

1. The outcome evaluation will be designed in its entirety by the academic research team so as to ensure it meets the expectation that a doctoral dissertation makes an original and substantive contribution to the field. The evaluation will be informed by discussions with key stakeholders at the Foundation for Youth Development. Stakeholders will include, but are not limited to, the following: the Manager of Programme Development, Ms.

Victoria Hood; the *Kiwi Can* Programme Developer, Ms. Brownyn Blair; the Manager of Research and Evaluation, Ms. Julie Moore and the *Kiwi Can* coordinators from each of the ten *Kiwi Can* regions.

2. Data collection tools used in the outcome evaluation process will reflect the requirements of the academic research staff and not, necessarily, the Foundation for Youth Development. The collection of this data will be used to inform Ms. Williams' PhD thesis and will be treated in an identical manner to all other data collected as part of the evaluation process. All collection, access and use of data will be subject to constraints imposed by the University of Auckland Human Participants Ethics Committee and applied equally to both parties.
3. The collection of data (eg. surveys) will be the shared responsibility of the academic research team, the Manager of Research and Evaluation at FYD, *Kiwi Can* coordinators, *Kiwi Can* leaders and/or teachers (where applicable).
4. The analysis of the quantitative and qualitative data gathered during the process of the evaluation will also be the primary responsibility of the academic research team. Analysis will be informed by discussions with the Manager of Research and Evaluation at FYD, *Kiwi Can* coordinators, *Kiwi Can* leaders and/or teachers (where applicable).
5. All data, whether it be quantitative or qualitative in nature, collected during the evaluation process will be jointly owned by the academic research team and the Foundation for Youth Development. The original hardcopies of each of the data collection tools will be located in an office on the Epsom Campus at the University of Auckland for the duration of Ms. Williams' PhD. Upon successful completion of the PhD the data will be stored for up to six years at a secure location approved by the University of Auckland. FYD retains the right to request copies of this data subject to UAHPEC restrictions.
6. Any electronic data collected or generated manually will be available to both the academic research team and the Foundation for Youth Development throughout the evaluation process. However, the Foundation for Youth Development must request permission from the academic research team to use the data in a public forum (e.g., presentations, marketing campaigns, etc.) prior to the successful completion of Ms. Williams' PhD. Any reasonable

request is not likely to be declined appropriately referenced and acknowledged.

7. Ms. Williams will have full rights to use the data collected during the outcome evaluation process to write her PhD thesis. The role of the FYD in the evaluation process will be fully acknowledged in her work.
8. Ms. Williams will have permission to access the Kiwi Can outcome evaluation database for five years from the date of her PhD defence.
9. Ms Williams will provide FYD with two written reports. A report on the implementation evaluation of Kiwi Can due June 2013, and a report on the outcome evaluation of Kiwi Can due June 2014, or on a date agreed to by both parties.
10. Ms Williams and FYD are committed to working in a collaborative manner to allow each party to achieve their objectives.
11. Ms. Williams will have permission to publish any findings from the Kiwi Can outcome evaluation in scholarly publications and/or presentations or talk about the research with the media or in any other public forum. The FYD will be given the right to comment and, where feasible, provided an advance copy of any dissemination. The FYD will comment in a timely fashion and will not impede the dissemination of research findings. All disseminations will keep confidential the identity of FYD participants, unless specifically requested otherwise.
12. There will be no automatic right to authorship for any publication and/or presentation simply arising from having participated in the data collection. Authorship will be granted according to the American Psychological Association standards and protocols. Where FYD is not involved as an author, authors will acknowledge the contribution of FYD.
13. FYD has the right to use any findings resulting from the research for marketing and fundraising purposes, to support their research and evaluation programme and for any other purpose not prohibited by this memorandum. The research team will be provided an advance copy, where feasible, of any dissemination and given the right to comment. The research team will comment in a timely fashion and will not impede the dissemination of public relations materials. All disseminations will keep confidential the identity of the research team, unless specifically requested otherwise.

14. 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.
15. Any official communication between the FYD, Kiwi Can, and the research team will be conducted via the FYD Manager of Research and Evaluation (Ms. Julie Moore) and/or the CEO (Ms. Marion Short) and Ms Williams on behalf of the research team.
16. Regardless of any personnel changes within the FYD or the research team at the University of Auckland during the time of this study, the research will continue according to the guidelines outlined in this document.
17. Ms Williams will provide a budget of expected outgoings associated with conducting the Kiwi Can outcome evaluation in conjunction with FYD.
18. Any expenses to be incurred by the University of Auckland research team in association with this project must be approved in writing by FYD.
19. FYD will reimburse the University of Auckland research team for actual and reasonable expenses, provided they have been authorized by FYD before being incurred. Reimbursement will be made on receipt of a tax invoice.

Signed by

Date

\_\_\_\_\_  
(Rachel Williams, PhD candidate)

\_\_\_\_\_  
(Dr. Gavin Brown, Supervisor)

\_\_\_\_\_  
(Representative of FYD)

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

## Appendix B



School of Learning, Development and Professional  
Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT ASSENT FORM

YEAR 6-8 STUDENTS

**THIS ASSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS**

**Title of Project:** Can it? Conducting an outcome evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand

**Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Supervisor) and Ms. Rachel Williams (University of Auckland PhD student)

I, ....., (print name) have read the Participation Information Sheet, and I understand what the research project is about.

I have been able to ask questions and to have them answered.

I understand that by signing this assent form, I will be participating in the research study.

I understand that I can request to leave the study at any time without giving a reason and that I can ask for my data to be withdrawn from the study by 31 December 2013.

I understand that the principal has made the assurance that the decision you make about participating in the study will have no impact on your learning, achievement or standing at school.

I understand that my personal information will remain confidential to the research team listed above.

I understand that my personal information and will not appear in any publications or reports that result from this research project.

I understand that I will not be identified in any publications or reports that arise from this research.

I understand that all data (paper copies) will be kept locked up for the duration of the research project. It will be stored either at The University of Auckland or in an approved long-term storage facility. It will be destroyed after six years. Any digital data files will be securely deleted from the university computer system.

I understand that the data from this research will be used in conference presentations, academic journal articles, a confidential report to The Foundation of Youth Development and in the student researcher's PhD.

Name: \_\_\_\_\_ (please print)

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number  
2013/8912

## Appendix C



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT INFORMATION SHEET

#### TEACHERS

- Title of Project:** A Survey of Health Education Delivery in Primary Schools in Auckland, New Zealand
- Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-Investigator) and Ms. Rachel Williams (University of Auckland PhD student)

#### Researcher Introduction

My name is Rachel Williams and I am working towards a PhD in the Faculty of Education at The University of Auckland. I am in my first year of study under the supervision of Associate Professor Gavin Brown and Dr. Pat Bullen. Prior to commencing my PhD I worked as a teacher and tutor in the United Kingdom, Europe, Japan and Canada. It was in Canada that I completed my Master of Public Health degree and became committed to studying issues related to health promotion and positive development in youth. As a potential participant in this study it is important that you read this document carefully.

#### Project description and procedures

The general aim of this research is to gather information on the implementation of the health education curriculum in Auckland-region primary schools. For this project, the opinions and personal experiences of all teachers from contributing and full-primary schools across Auckland will be sought. The research team is particularly interested in gathering information on the frequency with which teachers deliver lessons related to the strands, key learning areas and/or achievement objectives outlined in the health education curriculum. Information will be collected, analysed, and reported in three

different media - scholarly journal articles, research conference presentations, and/or the student researcher's PhD dissertation. The data for this project will be collected using an anonymous questionnaire. The questionnaire will be web-based (online) and it will take you approximately 15-20 minutes to complete. Overall, three thousand eight hundred teachers, like yourself, from all contributing and full-primary schools across Auckland will be invited to participate in the project. Information about the survey will be sent to you in electronic form. Hardcopies of the Participant Information Sheet have also been sent to your school. It is entirely up to you to decide whether or not you wish to participate in the survey. Further, you are free to withdraw from the study at any time without penalty, however, once you have submitted the anonymous questionnaire it will not be possible to remove your personal data from the study. Please note that we do not anticipate that you, as a participant, will experience any physical, psychological, or social discomfort as a result of participation. Lastly, you may choose to enter your email address at the end of the survey to be entered into a draw to win a brand new digital camera.

### **Data collection, analysis and retention**

The data from the online questionnaires will be entered into a computer-based statistical tool. It will be saved in a password-protected file on the personal computer of the Principal Investigator. Access to the electronic data will be restricted to the members of the research team. All data will be kept for six years to allow for publication and future re-analysis, after which time it will be shredded and deposited in a contracted, secure destruction system.

### **Anonymity and Confidentiality**

The privacy, anonymity and confidentiality of all participants in this study will be ensured using the anonymous Web Link collector tool available on Survey Monkey (the online survey tool used in this study). This tool assigns a number, or respondent ID, at random to each completed questionnaire. This random number will then be used during the analysis process to identify responses that were generated by a single respondent. Please note that participants are offered the opportunity to record their email address at the end of the survey in order to be entered into a draw to win a brand new digital camera. The email addresses will be used for the express purpose of

randomly selecting the winner of the digital camera. Once this process has been completed the email addresses will be permanently deleted from the questionnaires. Thank you very much for making this study possible. If you have any questions or concerns about the project please contact Ms. Rachel Williams (rwil313@aucklanduni.ac.nz; 623 8899 Ext 48384). For ethical concerns contact: The Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Private Bag 92019, Auckland, 1142. Telephone 09 373-7599 ext.87830/83761. Email: humanethics@auckland.ac.nz. Approved by The Auckland Human Participants Ethics Committee on 2 October 2012 for (3) years, Reference Number 2012/8362.

## Appendix D

### HEALTH AND PHYSICAL EDUCATION QUESTIONNAIRE - EXCERPT

This questionnaire is part of the third and final round of the large *National School Sampling Study*. In 2001 and 2002 several thousand teachers provided valuable information about their experiences in working from the *NZ Curriculum Framework* and one or more of the seven national curriculum statements. The findings, to date, are on the M.O.E. website: [www.moe.govt.nz](http://www.moe.govt.nz).

This questionnaire is an opportunity for you to provide information about your experiences in implementing the *Health and Physical Education in the New Zealand Curriculum* statement. The Ministry of Education is now seeking the views of a large sample of teachers, the results of which will influence future policy on curriculum and lead to sharing of practices between teachers and schools.

The information you provide will be seen only by the research team; individual schools and teachers will not be identifiable in reports of the results to the Ministry of Education. However, you may choose to provide your name for possible future internet or face-to-face discussions with other teachers about teaching experiences

In the questions we have used wording taken from the national curriculum:

- *The NZ Curriculum Framework* is the document published in 1993, which set out the overall policy direction for curriculum and assessment.
- By health and physical education **curriculum statement** we mean the statement released by the Ministry of Education called *Health and Physical Education in the New Zealand Curriculum (1999)*.
- The terms **strands**, **achievement objectives**, and **levels** refer to their meaning in the *NZ Curriculum Framework (1993)* and the seven national curriculum statements.

The focus of the questions is on teaching experiences with programmes related to the health and physical education curriculum statement (especially in years 1-10 where it is compulsory) so teachers of health and/or physical education, home economics, food and technology, and related areas are asked to respond; and so is not concerned with qualification changes such as the introduction of NCEA.



## IMPACT OF CURRICULUM STRUCTURE AND ASSESSMENT AND COMPLIANCE PRACTICES ON TEACHING AND LEARNING

We are interested in how you have responded to the way health and physical education as an area of learning is structured by the *New Zealand Curriculum Framework* and the health and physical education national curriculum statement, *Health and Physical Education in the New Zealand Curriculum*.

8. How user-friendly have you found the health and physical education curriculum statement?

Very user-friendly friendly	User-friendly	User-friendly in some ways	Not user-friendly
--------------------------------	---------------	----------------------------	-------------------

9. In your teaching programme related to health and physical education, to what extent has the health and physical education curriculum statement with its strands, levels, achievement objectives, underlying concepts etc. assisted you in:

- a) planning your school/department programme?

Always	Sometimes	Not very often	Rarely
	Don't know		

- b) planning your classroom programme?

Always	Sometimes	Not very often	Rarely
	Don't know		

- c) gaining an overview of the progression of the underlying concepts and skills?

Always	Sometimes	Not very often	Rarely
	Don't know		

- d) developing (specific) learning outcomes?

Always	Sometimes	Not very often	Rarely
	Don't know		

- e) achieving consistent understanding of the curriculum levels within your school?

Always	Sometimes	Not very often	Rarely
	Don't know		

- f) planning for assessment?

Always	Sometimes	Not very often	Rarely
	Don't know		

- g) meeting the needs of Māori students in your classroom/school?

Always	Sometimes	Not very often	Rarely
	Don't know		

h) communicating student achievement to other teachers within your school?

Always	Sometimes	Not very often	Rarely
	Don't know		

i) reporting student achievement to parents and caregivers?

Always	Sometimes	Not very often	Rarely
	Don't know		

j) communicating student achievement to teachers at other schools?

Always	Sometimes	Not very often	Rarely
	Don't know		

10. Would you make any changes to the **structure/organisation** of the health and physical education curriculum statement?

Yes

No

Don't know

If yes, what?

---

11. Would you make any changes to the **content** (strands, concepts etc.) of the health and physical education curriculum statement?

Yes

No

Don't know

If yes, what?

---

## STRANDS

12. How do you ensure coverage of the four strands - (i.e., **Personal Health and Physical Development; Movement Concepts and Motor Skills; Relationships with Other People; and Healthy Communities and Environments**) - in the health and physical education curriculum statement? (*For example, do you teach all strands in each unit? Do you teach the strands in an integrated way?*)

---

---

---

13. Overall I give:

- more or less equal emphasis to each strand.
- more emphasis to Strand A – Personal Growth and Development.
- more emphasis to Strand B – Movement Concepts and Motor Skills.
- more emphasis to Strand C – Relationships with Other People.
- more emphasis to Strand D – Healthy Communities and Environments.

## KEY AREAS OF LEARNING

14. What key learning areas are taught **in your school**? (NB. not necessarily taught separately; *please tick appropriate box/es*)
- Body Care and Physical Safety
  - Food and Nutrition
  - Mental Health
  - Outdoor Education
  - Physical Activity
  - Sexuality Education
  - Sports Studies
  - Other** (*please specify*):
15. What key learning areas **do you teach**? (*please tick appropriate box/es*)
- Body Care and Physical Safety
  - Food and Nutrition
  - Mental Health
  - Outdoor Education
  - Physical Activity
  - Sexuality Education
  - Sports Studies
  - Other** (*please specify*)

## Appendix E

### Health Education Questionnaire (Study #1)

#### Health Education Practice in New Zealand Primary Schools: A Questionnaire for Teachers

This questionnaire is part of a PhD study on health education theory, design and practice undertaken by Rachel Williams, a PhD student in the Faculty of Education at The University of Auckland. It is intended to provide evidence about the current state of health education delivery in Auckland-region primary schools.

The information that you provide will be seen only by the research team (i.e., Rachel and her 2 supervisors—Assoc. Prof. Gavin Brown and Dr. Pat Bullen); individual schools and teachers will not be identifiable in any written or oral reports of these findings.

In the questionnaire all wording has been taken from the two primary health curriculum documents in New Zealand:

- The *Health and Physical Education in the New Zealand Curriculum* statement released by the Ministry of Education in 1999; and
- *The New Zealand Curriculum* document released by the Ministry of Education in 2007

The terms **strands**, **key learning areas**, and **levels** refer to their meaning in *The New Zealand Curriculum* document (2007).

## INTRODUCTION

1. What is your position in your school? *(please tick appropriate boxes)*
  - Principal
  - Teaching Principal
  - Deputy/Assistant/Associate Principal
  - Teaching Deputy/Asst/Associate Principal
  - Curriculum/Syndicate Leader
  - Head of Department/Head of Faculty
  - Classroom Teacher/Subject Teacher
  - Other (please specify):** \_\_\_\_\_
  
2. In which region of Auckland is your school located? *(please tick appropriate box)*
  - Auckland Central
  - Manukau
  - North Shore
  - Rodney
  - Waitakere
  - Papakura
  - Waiheke Island
  - Franklin
  
3. How many years have you been teaching? *(include both part-time and full-time)*
  - Less than (<) 2 years
  - 2-5 years
  - 6-10 years
  - 11-15 years
  - More than (>) 15 years
  
4. How many years have you been teaching programmes that relate to **Health Education** in *The New Zealand Curriculum*? *(include both part-time and full-time)*
  - < 2 years
  - 2-5 years
  - 6-10 years
  - 11-15 years
  - > 15 years
  
5. What class level(s) do you currently teach programmes related to **health education**? *(please tick appropriate boxes)*
  - Year 1
  - Year 2
  - Year 3
  - Year 4
  - Year 5
  - Year 6
  - Year 7
  - Year 8

## THE HEALTH EDUCATION CURRICULUM – DIMENSIONS OF HEALTH

In the New Zealand curriculum framework and document (1999 and 2007) there are four concepts that underlie the Health and Physical Education learning area. They are *Hauora*, *Attitudes and Values*, *the Socio-ecological perspective* and *Health Promotion*.

In the curriculum, *Hauora* (or well-being), is depicted using the Te Whare Tapa Wha Model (see below).

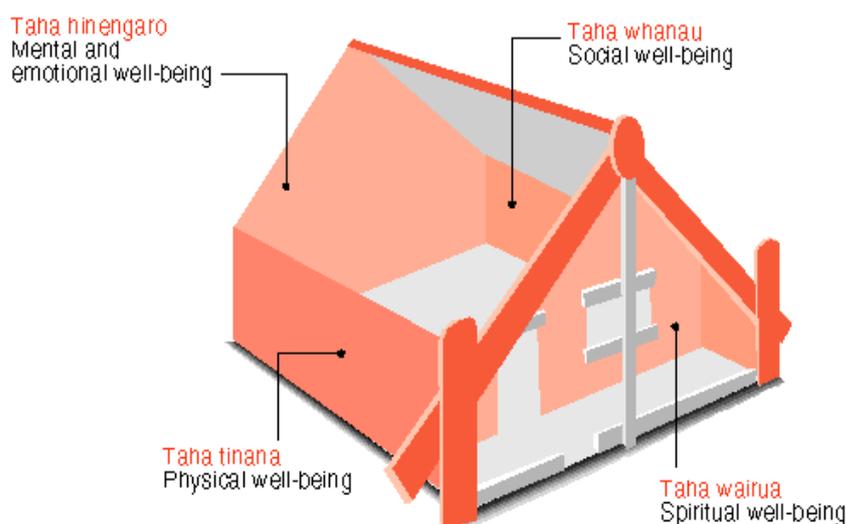


Figure 1.0 Te Whare Tapa Wha Model (Durie, 1999)

6. In your opinion, how important is it to teach your students about the five dimension(s) of health?

Dimension of health	Very important	Important	Moderately important	Slightly important	Not important
Physical health	<input type="checkbox"/>				
Emotional health	<input type="checkbox"/>				
Mental health	<input type="checkbox"/>				
Spiritual health	<input type="checkbox"/>				
Social health	<input type="checkbox"/>				

7. In your opinion, how influential is each dimension of health on student well-being?

Dimension of health	Very influential	Influential	Moderately influential	Slightly influential	Not influential
Physical health	<input type="checkbox"/>				
Emotional health	<input type="checkbox"/>				
Mental health	<input type="checkbox"/>				
Spiritual health	<input type="checkbox"/>				
Social health	<input type="checkbox"/>				

8. In your personal practice, how often do you teach students about each of the dimensions of health?

Dimension of health	Always/ Very Frequently (In every or nearly every lesson)	Frequently (About ¾ to ½ of lessons)	Occasionally (Around 1/3 of lessons)	Seldom (In less than 1/4 of the lessons)	Very rarely/ Never (No more than once)
Physical health	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Emotional health	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mental health	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Social health	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Spiritual health	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. How satisfied are you with the amount of time that you dedicate to teaching each dimension?

Dimension of health	Very satisfied	Satisfied	Moderately satisfied	Slightly satisfied	Dissatisfied
Physical health	<input type="checkbox"/>				
Emotional health	<input type="checkbox"/>				
Mental health	<input type="checkbox"/>				
Social health	<input type="checkbox"/>				
Spiritual health	<input type="checkbox"/>				

### THE HEALTH EDUCATION CURRICULUM – GENERAL IMPLEMENTATION

10. Over the course of the year, how do you deliver health education classes? (*please tick appropriate box/es*)

- In blocks/modules
- As a subject with its own timetable slots
- As a stand-alone subject, but with teachers with various backgrounds planning and delivering different parts
- Integrated into other learning areas: (*please tick appropriate boxes below*)

English  
Languages  
Science  
Te Reo Maori

Mathematics  
Social Studies  
The Arts

Pacific and International  
Technology  
Other (*please specify*):

- I do not deliver the health education curriculum. It is delivered by an external provider. (If you ticked this box please move to Question #18)

*For the following questions, please refer only to those classes that are 'stand-alone' health education classes (ie. Lessons that are taught by you, dedicated to health education and not integrated into other learning areas)*

11. On average, how many health education classes do you teach over the course of a week? *(please fill in the number in the space provided)*

\_\_\_\_\_ classes per week

12. On average, how long is each health education lesson? *(please fill in the number of minutes in the space provided)*

\_\_\_\_\_ minutes per lesson

**Key Learning Areas**

13. How frequently are the learning areas in health education taught in your school? *(please tick appropriate boxes)*

Strand/Frequency	Always/Very frequently	Frequently	Occasionally	Seldom	Very rarely/Never
Body Care and Physical Safety	<input type="checkbox"/>				
Food and Nutrition	<input type="checkbox"/>				
Mental Health	<input type="checkbox"/>				
Sexuality Education	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

14. How frequently do you teach the key learning areas? *(please tick appropriate boxes)*

Strand/Frequency	Always/Very frequently	Frequently	Occasionally	Seldom	Very rarely/Never
Body Care and Physical Safety	<input type="checkbox"/>				
Food and Nutrition	<input type="checkbox"/>				
Mental Health	<input type="checkbox"/>				
Sexuality Education	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

**Achievement Objectives (AO)**

15. How frequently do you address the following achievement objective categories from The New Zealand Curriculum (2007) in your health education classes? *(please tick all appropriate boxes)*

AO/Frequency	Always/Very frequently	Frequently	Occasionally	Seldom	Very rarely/Never
Personal growth and development	<input type="checkbox"/>				
Safety Management	<input type="checkbox"/>				
Personal identity	<input type="checkbox"/>				
Positive attitudes	<input type="checkbox"/>				
Relationships	<input type="checkbox"/>				
Identity,	<input type="checkbox"/>				

sensitivity and respect					
Interpersonal skills	<input type="checkbox"/>				
Societal attitudes and values	<input type="checkbox"/>				
Community resources	<input type="checkbox"/>				
Rights, responsibilities and laws	<input type="checkbox"/>				
People and the environment	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

### Strands

16. The strands most relevant to health education in The New Zealand Curriculum document (2007) are: *Personal health and Physical Development, Relationships with Other People* and *Healthy Communities and Environments*. Overall, how much emphasis do you give to these strands in your lessons? Rank order the strands from 1 (most emphasis) to 3 (least emphasis).

- \_\_\_\_\_ Strand A – Personal Growth and Development  
 \_\_\_\_\_ Strand C – Relationships with Other People  
 \_\_\_\_\_ Strand D – Healthy Communities and Environments

### Approach

17. Are your health education classes primarily knowledge-based (ie. Involve educating students about health facts) or skill-based (ie. Focus on developing specific health-related skills)? (please indicate your response by placing an 'X' in the appropriate location on the continuum)

\_\_\_\_\_

Knowledge-based Skill-based

18. In your opinion, should the focus of health education classes be more knowledge-based or more skill-based? (please indicate your opinion by placing an X in the appropriate location on the continuum)

\_\_\_\_\_

Knowledge-based Skill-based

## BARRIERS TO DELIVERY

19. How strongly have these potential barriers affected your ability to implement the health education curriculum? *(please tick appropriate boxes)*

Barrier	Very important	Important	Moderately important	Slightly important	Not important
Lack of time	<input type="checkbox"/>				
Lack of resources	<input type="checkbox"/>				
Lack of personal knowledge in the subject area	<input type="checkbox"/>				
Lack of interest from students	<input type="checkbox"/>				
Lack of personal interest in the subject area	<input type="checkbox"/>				
Lack of support from school management	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

## PERSONAL COMMENTS

20. Do you have any additional comments that you would like to make about the delivery of health education in your classroom and/or school?

If you are interested in having your email address entered into a draw to win one of two \$50.00 Westfield gift vouchers, please enter it below:

*Thank you very much for your participation. It is greatly appreciated!*

# Appendix F

Kiwi Can and the New Zealand Curriculum							
Themes		KEY Heart Values Mind Values					
<b>Positive Relationships</b> <i>Whakawhanaukatanga</i> <i>...working and playing in safe and thoughtful ways with others.</i>	<b>Integrity</b> <i>Pono</i> <i>...doing the right thing.</i>	<b>Resilience</b> <i>Taikaha</i> <i>...coping with challenges and 'bouncing back.'</i>	<b>Respect</b> <i>Manaakitanga</i> <i>...treating people, things and places as important.</i>				
				<b>Kiwi Can Modules: (Set A)</b> Positive Communication Co-operation Friendship	<b>Kiwi Can Modules: (Set A)</b> Honesty Responsibility Reliability	<b>Kiwi Can Modules: (Set A)</b> Understanding Emotions Self control / Self discipline Dealing with Challenges	<b>Kiwi Can Modules:</b> For Ourselves**
				<b>Relationships with Other People</b> <b>Relationships</b> L1 Explore and share ideas about relationships with other people. L2 Identify and demonstrate ways of maintaining and enhancing relationships between individuals and within groups. L3 Identify and compare ways of establishing relationships and managing changing relationships. L4 Identify the effects of changing situations, roles, and responsibilities on relationships, and describe appropriate responses.	<b>Relationships with Other People</b> <b>Relationships</b> L1 Explore and share ideas about relationships with other people. L2 Identify and demonstrate ways of maintaining and enhancing relationships between individuals and within groups. L3 Identify and compare ways of establishing relationships and managing changing relationships. L4 Identify the effects of changing situations, roles, and responsibilities on relationships and describe appropriate responses.	<b>Relationships with Other People</b> <b>Interpersonal Skills</b> L1 Express their own ideas, needs wants and feelings clearly, and listen to those of other people. L2 Express their ideas needs wants and feelings appropriately and listen sensitively to other people and affirm them. L3 Identify the pressures that can influence interactions and other people and demonstrate basic assertiveness strategies to manage these. L4 Describe and demonstrate a range of assertive communication skills and processes that enable them to interact appropriately with other people.	<b>Personal Health and Development:</b> <b>Personal Identity</b> L1 Describe themselves in relation to a range of contexts. L2 Identify personal qualities that contribute to a sense of self-worth. L3 Describe how their own feelings, beliefs, and actions, and those of other people, contribute to their personal sense of self-worth. L4 Describe how social messages and stereotypes, including those in the media, can affect feelings of self-worth.
				<b>Kiwi Can Modules: (Set B)</b> Fairness and Fairplay Leadership Conflict Resolution	<b>Kiwi Can Modules: (Set B)</b> Making Good Choices Being a Role Model Accountability	<b>Kiwi Can Modules: (Set B)</b> Goal Setting Problem Solving Perseverance	<b>Kiwi Can Modules:</b> Respectful Communication ** For Others** For our School** For our Community** For our Environment**
<b>Relationships with Other People:</b> <b>Interpersonal Skills</b> L1 Express their own ideas, needs wants and feelings clearly, and listen to those of other people. L2 Express their ideas needs wants and feelings appropriately and listen sensitively to other people and affirm them. L3 Identify the pressures that can influence interactions and other people and demonstrate basic assertiveness strategies to manage these. L4 Describe and demonstrate a range of assertive communication skills and processes that enable them to interact appropriately with other people.	<b>Personal Health and Development:</b> <b>Personal Identity</b> L1 Describe themselves in relation to a range of contexts. L2 Identify personal qualities that contribute to a sense of self-worth. L3 Describe how their own feelings, beliefs, and actions, and those of other people, contribute to their personal sense of self-worth. L4 Describe how social messages and stereotypes, including those in the media, can affect feelings of self-worth.	<b>Personal Health and Development:</b> <b>Personal Identity</b> L1 Describe themselves in relation to a range of contexts. L2 Identify personal qualities that contribute to a sense of self-worth. L3 Describe how their own feelings, beliefs, and actions, and those of other people, contribute to their personal sense of self-worth. L4 Describe how social messages and stereotypes, including those in the media, can affect feelings of self-worth.	<b>Relationships with Other People:</b> <b>Relationships</b> L1 Explore and share ideas about relationships with other people. L2 Identify and demonstrate ways of maintaining and enhancing relationships between individuals and within groups. L3 Identify and compare ways of establishing relationships and managing changing relationships. L4 Identify the effects of changing situations, roles and responsibilities on relationships and describe appropriate responses.				
<b>School Community Project</b>		<b>Healthy Communities and Environments: Rights Responsibilities and Laws: People and Environment</b> L1 Take individual and collective action to contribute to environments that can be enjoyed by all. L2 Contribute to and use simple guidelines and practices that promote physically and socially healthy classrooms, schools and local environments. L3 Plan and implement a programme to enhance an identified social and physical aspect of their classroom or school environment. L4 Specify individual responsibilities and take collective action for the care and safety of other people in their school and in the wider community.					
<b>Key Competencies</b>		<b>Thinking, Managing Self, Participating and Contributing, Relating to Others, The Use of Language Symbols and Texts</b>					

Note. Reprinted with permission for the Foundation for Youth Development.

## Appendix G



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT INFORMATION SHEET

#### YEAR 4-6 STUDENT

- Title of Project:** Can it? An implementation evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand
- Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-Investigator) and Ms. Rachel Williams (University of Auckland PhD student)

#### Researcher Introduction

My name is Rachel Williams and I am a PhD student at The University of Auckland. I am in my first year of study. My supervisors are Associate Professor Gavin Brown and Dr. Pat Bullen. For my PhD I want to study health and well-being in children. You have been given this sheet because I would like you to be in my study. Please read the whole sheet. It will tell you about my research project. You do not have to participate in this study. It is your choice.

#### The Research Project

The goal of this research is to gather information about the *Kiwi Can* lessons. I want to know about your experiences of the *Kiwi Can* programme. If you choose to participate, you will complete one questionnaire. You will not have to write your name on the questionnaire. It is anonymous. The questionnaire will take about 20 minutes to complete. You will fill it out during one of your *Kiwi Can* lessons.

I will read the whole questionnaire out loud to you. If you have any questions you can ask me. If you want to stop participating at any time you can. You can put your pencil

down and quietly leave the *Kiwi Can* classroom and go into the hallway. Your *Kiwi Can* leaders will sit with you in the hallway until the study has finished.

If you do not want to participate in the study at all that is fine. No one, including the researchers, your parents, your friends and your teachers, will know if you have participated or not. Your decision will not affect your learning or your grades at school.

You have been chosen for this study because you go to a *Kiwi Can* school and you are in Year 4, Year 5, or Year 6. You are one of three hundred students who have been asked to participate.

All of the completed questionnaires will be given to me. I will take them back to the University of Auckland and look at them. I will then write a report, a journal article and a paper (thesis) for my school. I may also present my findings at a conference. I will never use your name in any of my work/publications.

All of the questionnaires will be locked in a cabinet at the university or in a secure storage place. The questionnaires will be locked away for six years and then they will be put into a secure destruction system.

Thank you for reading about my study!

If you have any questions about the project please contact me:

Rachel Williams

[rwil313@aucklanduni.ac.nz](mailto:rwil313@aucklanduni.ac.nz)

623 8899 Ext 48384

For ethical concerns contact: The Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Private Bag 92019, Auckland, 1142. Telephone 09 373-7599 ext.87830/83761. Email: [humanethics@auckland.ac.nz](mailto:humanethics@auckland.ac.nz)

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 01 JUNE 2012 for (3) years, Reference Number 2012/8039

## Appendix H



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT **NON**-CONSENT FORM

PARENTS/CAREGIVERS

**THIS **NON**-CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX  
YEARS**

**Title of Project:** Can it? Conducting an implementation evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand

**Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Supervisor) and Ms. Rachel Williams (University of Auckland PhD student)

I, ....., (print name) have read the Participation Information Sheet, have understood the nature of the research and why my son/daughter has been selected to participate.

I have had the chance to ask questions and to have them answered to my satisfaction.

After reading through the Participant Information Sheet I do **NOT** agree that my son/daughter (circle one) can participate in this research.

I understand that by signing this consent form, my son/daughter will **NOT** be participating in the research study.

I understand that my son/daughter will be given an alternative task to complete while his/her classmates complete the questionnaire.

Name: \_\_\_\_\_

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 01 JUNE 2012 for (3) years, Reference Number  
2012/8039

## Appendix I



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### Kiwi Can Programme Questionnaire (IE)

Today's date: \_\_\_\_\_

#### Part A. Personal Information

Please read the following questions and answer them by shading in the right box.

Example: Do you like ice cream?

- Yes  
 No

1. Are you a boy or a girl?

- Boy  
 Girl

2. How old are you?

- 7 years old  
 8 years old  
 9 years old  
 10 years old  
 11 years old  
 12 years old

3. What is your Year Level?

- Year 4  
 Year 5  
 Year 6

4. How long have you been at this school?

- 1 year  
 2 years  
 3 years  
 4 years  
 5 years  
 6 years  
 7 years  
 8 years

5. Which people group(s) do you belong to? (You can shade more than one box).

- Māori
- Pasifika
- Pakeha/New Zealand European
- Asian
- Other

**Part B. The Kiwi Can Programme Lessons**

Which of the following things have you done in your Kiwi Can classes this year?

I remember learning about being:

- Respectful
- Honest
- Responsible
- Motivated
- Able to cope with challenges
- Polite
- A good friend
- A good team member

I remember:

- Making goals
- Solving problems
- Talking nicely to other people
- Being kind to others
- Co-operating with others
- Taking turns

**Part C.** Please read each sentence below. How strongly do you agree or disagree with the following statements about Kiwi Can lessons?

		Disagree	Slightly Disagree	Slightly Agree	Mostly agree	Agree
1	Overall, I really like the Kiwi Can lessons.	<input type="checkbox"/>				
2	I am often bored during Kiwi Can lessons.	<input type="checkbox"/>				
3	I participate in all aspects of the Kiwi Can lesson.	<input type="checkbox"/>				
4	Kiwi Can lessons are the best part of the school week.	<input type="checkbox"/>				
5	I have learned new things from taking part in the Kiwi Can lessons.	<input type="checkbox"/>				
6	Kiwi Can lessons teach me important skills.	<input type="checkbox"/>				
7	Kiwi Can lessons teach me about values.	<input type="checkbox"/>				
8	The Kiwi Can lessons follow the same pattern (Recap, Energiser, Catch Phrase and Activity)	<input type="checkbox"/>				
9	We have a topic in each Kiwi Can lesson.	<input type="checkbox"/>				
10	My Kiwi Can lessons start on time.	<input type="checkbox"/>				
11	My Kiwi Can lessons are well organized.	<input type="checkbox"/>				
12	We use equipment in our Kiwi Can lessons.	<input type="checkbox"/>				
13	There is enough equipment for everyone in the class.	<input type="checkbox"/>				
14	I am taught how to use the Kiwi Can equipment safely.	<input type="checkbox"/>				
15	I look up to my Kiwi Can leaders.	<input type="checkbox"/>				
16	My Kiwi Can leaders are energetic.	<input type="checkbox"/>				

**THANK YOU FOR YOUR PARTICIPATION!**

## Appendix J



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### Measure of Student Social Health (KCOE)

#### Part A. Personal Information

Please read the following questions and answer them by shading in the right box.

Example: Do you like ice cream?

- Yes  
 No

1. Are you a boy or a girl?

- Boy  
 Girl

2. What is your age?

- 10 years of age  
 11 years of age  
 12 years of age  
 13 years of age  
 14 years of age

3. What is your Year Level?

- Year 6  
 Year 7  
 Year 8

4. How long have you been at this school?

- 1 year  
 2 years  
 3 years  
 4 years  
 5 years  
 6 years  
 7 years  
 8 years  
 9 years

5. Which people group(s) do you belong to? (You can shade more than more box).

- Maori  
 Pasifika  
 Pakeha/New Zealand European  
 Asian  
 Other

6. Is Kiwi Can in your school?

- Yes  
 No – If you answered “no” please move to Part B

7. Do you go to Kiwi Can classes each week?

- Yes  
 No

**Part B. All About You**

Please read each sentence below. How strongly do you agree or disagree with the following statements?

		Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
1	I often feel sorry for other children who are sad or are in trouble	<input type="checkbox"/>					
2	I believe that there are two sides to every question and try to look at them both	<input type="checkbox"/>					
3	I say thank you when someone does something nice for me	<input type="checkbox"/>					
4	Others like me and have fun with me	<input type="checkbox"/>					
5	I kick or hit someone else when they make me angry	<input type="checkbox"/>					

		Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
6	I care about the school I go to	<input type="checkbox"/>					
7	My classmates are there when I need them	<input type="checkbox"/>					
8	I respect the values and beliefs of people who are of a different culture than I am	<input type="checkbox"/>					
9	I stand up for what I believe in, even when it's unpopular to do so	<input type="checkbox"/>					
10	I tell the truth even when it's not easy	<input type="checkbox"/>					

## Part B. All About You

Please read each sentence below. How strongly do you agree or disagree with the following statements?

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
11	I know a lot about people from other cultures	<input type="checkbox"/>					
12	My classmates really care about me	<input type="checkbox"/>					
13	My teachers really care about me	<input type="checkbox"/>					
14	I am bossy	<input type="checkbox"/>					
15	I make friends easily	<input type="checkbox"/>					

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
16	I take turns with others	<input type="checkbox"/>					
17	I sometimes try to understand my friends better by imagining how things look from their point of view	<input type="checkbox"/>					
18	I often feel sorry for people who don't have the things that I have	<input type="checkbox"/>					
19	When I see someone being picked on, I feel upset for them	<input type="checkbox"/>					
20	Before criticising someone I try to imagine how I would feel if I were being criticised	<input type="checkbox"/>					

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
21	I listen to others when they talk	<input type="checkbox"/>					
22	Others do not like me	<input type="checkbox"/>					
23	I speak or interrupt if someone else is talking	<input type="checkbox"/>					
24	Students in my school care about me	<input type="checkbox"/>					
25	I feel my classmates are my friends	<input type="checkbox"/>					

**Part B. All About You**

Please read each sentence below. How strongly do you agree or disagree with the following statements?

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
26	I enjoy being with people who are from a different culture than I am	<input type="checkbox"/>					
27	I accept responsibility for my actions when I make a mistake or get into trouble	<input type="checkbox"/>					
28	I do what I believe is right even if my friends make fun of me	<input type="checkbox"/>					
29	It is important to me to get to know people who are of a different culture than I am	<input type="checkbox"/>					
30	I trust my classmates	<input type="checkbox"/>					

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
31	I get a lot of encouragement at my school	<input type="checkbox"/>					
32	I share games and toys with others	<input type="checkbox"/>					
33	When I am upset with someone I usually try to “put myself in their shoes” for a while	<input type="checkbox"/>					
34	It makes me sad to see someone who can’t find anyone to play with	<input type="checkbox"/>					
35	I get upset when I see someone being hurt	<input type="checkbox"/>					

### Part B. All About You

Please read each sentence below. How strongly do you agree or disagree with the following statements?

		Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
36	I try to look at everybody's side of a disagreement before I make a decision	<input type="checkbox"/>					
37	I say I'm sorry when I hurt someone by accident	<input type="checkbox"/>					
38	I get along well with my teachers	<input type="checkbox"/>					
39	I get along well with my classmates	<input type="checkbox"/>					
40	I do my best even when I have to do a job I don't like	<input type="checkbox"/>					

### Part C. All About Your Class/School

Please read each sentence below. How strongly do you agree or disagree with the following statements about students in your school?

		Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
1	Students in my class play fair during games	<input type="checkbox"/>					
2	Students in my class borrow things without asking	<input type="checkbox"/>					
3	Students in my class help each other learn	<input type="checkbox"/>					
4	Students in my school feel safe in all areas of the school building	<input type="checkbox"/>					
5	My classroom is a fun place to be	<input type="checkbox"/>					

		Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
6	I am glad to get back to school after summer holiday	<input type="checkbox"/>					
7	Students in my school feel safe on the playground	<input type="checkbox"/>					
8	Students in my class work together to solve problems	<input type="checkbox"/>					
9	Students in my class laugh at other student's mistakes	<input type="checkbox"/>					
10	Students in my class cheer up someone who is feeling sad	<input type="checkbox"/>					

**Part C. All About Your Class/School**

Please read each sentence below. How strongly do you agree or disagree with the following statements about students in your school?

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
11	Students in my class keep promises that they make to other students	<input type="checkbox"/>					
12	Students in my class make fun of other students	<input type="checkbox"/>					
13	Students at this school really care about each other	<input type="checkbox"/>					
14	Students in my school feel safe on the school field	<input type="checkbox"/>					
15	I like my school	<input type="checkbox"/>					

		<b>Strongly Disagree</b>	<b>Mostly Disagree</b>	<b>Slightly Disagree</b>	<b>Slightly Agree</b>	<b>Mostly Agree</b>	<b>Strongly Agree</b>
16	I would be very sad if I had to go to a different school	<input type="checkbox"/>					
17	Students in this school treat each other with respect	<input type="checkbox"/>					
18	Students in my class bother other students when they are working	<input type="checkbox"/>					
19	Students in my class share with other students	<input type="checkbox"/>					
20	Students in my class follow the teacher's rules	<input type="checkbox"/>					
21	Students in my class talk without raising their hands during classroom discussions	<input type="checkbox"/>					
22	Students at this school are willing to go out of their way to help someone	<input type="checkbox"/>					

**THANK YOU FOR YOUR PARTICIPATION!**

## Appendix K



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT INFORMATION SHEET

#### PARENT/CAREGIVER – KIWI CAN SCHOOL

- Title of Project:** Can it? An outcome evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand
- Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-investigator) and Ms. Rachel Williams (University of Auckland PhD student)

#### Researcher Introduction

My name is Rachel Williams and I am working towards a PhD in the Faculty of Education at The University of Auckland. I am in my second year of study under the supervision of Associate Professor Gavin Brown and Dr. Pat Bullen. Prior to commencing my PhD I worked as a teacher and tutor in the United Kingdom, Europe, Japan, and Canada. It was in Canada that I completed my Master of Public Health degree and became committed to studying issues related to health promotion and positive development in youth.

As parents/caregivers of potential participants in the following study, it is important that you read this document carefully so that you can make an informed decision about whether you agree to have your son/daughter participate.

## **Project description and procedures**

The general aim of this project, an outcome evaluation of *Kiwi Can*, is to evaluate the impact that participation in the programme may be having on student attitudes, interpersonal (social) behaviours and/or the social climate of the school. The data will be collected using questionnaires.

In this study, a convenience sample of approximately 1200 Year 6-8 students from eighteen low-decile, intermediate and/or full-primary schools from around New Zealand will be selected according to a stratified sampling frame. Six of the schools will be 'new' *Kiwi Can* schools, six will be 'experienced' *Kiwi Can* schools and six will be 'control' (non-*Kiwi Can*) schools. 'New' *Kiwi Can* schools are defined as those that have been in the programme for less than two years while 'experienced' schools have been involved for three years or more. 'Control' schools are those who have never participated in the *Kiwi Can* programme. Once the schools have been selected, all Year 6-8 students in those schools will be invited to complete the same questionnaire. They will be asked to fill it out twice during the 2013 school year. The first period of data collection will be during Weeks 2 and 3 of Term I (February 2013). The second period of data collection will be during Weeks 6 and 7 of Term IV (November 2013). In *Kiwi Can* schools, like the one your child attends, the *Kiwi Can* leaders will administer the questionnaire during a single *Kiwi Can* lesson time in February and again in November. In control schools, the classroom teacher will be asked to administer the questionnaire. Before administering the questionnaire in your school the leaders will be fully trained by the student researcher. They will also be asked to sign a confidentiality agreement to indicate that they won't view nor share the student's responses on the questionnaire with others. All *Kiwi Can* leaders will be asked to follow a written script and to read the entire questionnaire aloud to the students.

The questionnaire itself is divided into three sections. In the first section students are asked for basic demographic information (Part A). This includes their name. In order to protect their identity, the research team will remove the student's name and replace it with a number between 1 and 1200 before the data is analysed or published. We will keep a master list of student names and corresponding questionnaire numbers in a password protected file on the personal computer of the student researcher. In the second and third sections of the questionnaire, students will be asked to read a series of

statements related to personal attitude and behavior (Part B) and school climate (Part C). They will each be given a 25-30 minutes to complete the questionnaire and leaders/teachers will give the students a five-minute break between Parts B and C.

I, the researcher, (Rachel Williams) will travel to each participating school to deliver the questionnaires, pens, Parent and Student Participant Information Sheets, Parent Non-Consent Forms and Student Assent Forms to the leaders one week prior to administration. When I am there I will be happy to answer any questions that you or your child may have about the research study.

Please note that there is no financial remuneration being offered for participating in the project. Further, as parents of the student participants you are asked to complete a Parental Non-Consent Form if you do not wish for your son/daughter to participate in the research project. Please return the Non-Consent Form to the school. Students who are not granted permission to participate in the study will be given an alternative task to complete. All students are under no obligation to participate and are free to withdraw their participation at any time without giving a reason and without negative consequences. The participants may also withdraw data that they have submitted until 31 December 2013. Please note that we do not anticipate that participants will experience any physical, psychological or social discomfort as a result of participation. Lastly, the school principal has made the assurance that the students' decision whether or not to participate in the study will have no impact on their learning, achievement or standing at school.

### **Data collection, analysis and retention**

The data from the questionnaires will be entered into a computer-based statistical tool. The hardcopies of the questionnaires will then be stored in a locked cabinet on The University of Auckland (Epsom Campus) premises or in an approved long-term storage facility. The Consent Forms will also be stored in a separate, secure location on the same premises. Access to the Consent Forms will be restricted to the research team directly involved in this project. The data will be jointly owned by the research team and The Foundation for Youth Development.

All data will be kept for six years to allow for publication and future re-analysis, after which time the hardcopy data will be deposited in a contracted, secure destruction system and the digital data will be securely deleted from the university computer system.

### **Anonymity and Confidentiality**

The privacy and confidentiality of all participants in this study will be ensured. For ethical concerns contact: The Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Private Bag 92019, Auckland, 1142. Telephone 09 373-7599 ext.87830/83761. Email: [humanethics@auckland.ac.nz](mailto:humanethics@auckland.ac.nz)

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number 2013/8912

## Appendix L



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT INFORMATION SHEET

#### YEAR 6-8 STUDENT – KIWI CAN SCHOOL

- Title of Project:** Can it? An outcome evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand
- Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-Investigator) and Ms. Rachel Williams (University of Auckland PhD student)

#### Researcher Introduction

My name is Rachel Williams and I am a PhD student at The University of Auckland. I am in my second year of study. My supervisors are Associate Professor Gavin Brown and Dr. Pat Bullen. For my PhD I want to study health and well-being in children. You have been given this sheet because I would like you to be in my study. Please read the whole sheet. It will tell you about my research project. You do not have to participate in this study. It is your choice.

#### The Research Project

The goal of this research is to gather information about the impact that *Kiwi Can* has on your attitude and behavior as well as the attitudes and behaviours of other students in your school. If you choose to participate, you will complete the same questionnaire at two different points during 2013. The questionnaire will take about 25-30 minutes to complete. You will fill it out during one of your *Kiwi Can* lessons.

Your *Kiwi Can* leaders will read the whole questionnaire out loud to you. If you have any questions you can ask them at any time. If you want to stop participating at any time you can. You can put your pencil down and quietly leave the *Kiwi Can*

classroom and go into the hallway. Your *Kiwi Can* leaders will sit with you in the hallway until the study has finished. Please note that your principal has made the assurance that the decision you make about participating in the study will have no impact on your learning, achievement or standing at school.

If you do not want to participate in the study at all that is fine. No one, including the researchers, your parents, your friends and your teachers, will know if you have participated or not. If you choose not to participate you will be asked to sit quietly while your *Kiwi Can* leader reads the questionnaire aloud to the class. Please note that your decision will not affect your learning or your grades at school. You can also ask to have your data withdrawn from the study until 31 December 2013. After that date you cannot ask for it to be withdrawn.

You have been chosen for this study because you go to a *Kiwi Can* school and you are in Year 6, Year 7, or Year 8. You are one of twelve hundred students who have been asked to participate.

All of the completed questionnaires will be mailed to me at The University of Auckland. I will immediately replace your name on the questionnaire with a number between 1 and 1200. At the same time I will create a list of names and numbers that is stored on my personal computer. No one else can see it and no one other than me will ever know what your personal responses were. I will use the answers on the questionnaire to write two reports, a journal article and a paper (thesis) for my school. I may also present my findings at a conference. I will never use your name in any of my work/publications.

All of the questionnaires will be locked in a cabinet at the university or in a secure storage place. The questionnaires will be locked away for six years and then they will be put into a secure destruction system. The digital data will be securely deleted from the university computer system.

Thank you for reading about my study!

If you have any questions about the project please contact the research team:

Rachel Williams [rwil313@aucklanduni.ac.nz](mailto:rwil313@aucklanduni.ac.nz); 623 8899 Ext 48384.

For ethical concerns contact: The Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Private Bag 92019, Auckland, 1142. Telephone 09 373-7599 ext.87830/83761. Email: [humanethics@auckland.ac.nz](mailto:humanethics@auckland.ac.nz)

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number 2013/8912

**PARTICIPANT CONSENT FORM**

**PARENTS/CAREGIVERS**

**THIS NON-CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX  
YEARS**

**Title of Project:** Can it? Conducting an outcome evaluation of *Kiwi Can*,  
a school-based values and life-skills programme, in  
Aotearoa New Zealand

**Research Team:** Associate Professor Gavin Brown (Principal  
Investigator), Dr. Pat Bullen (Supervisor) and Ms.  
Rachel Williams (University of Auckland PhD student)

I, ....., (print name) have read the  
Participation Information Sheet, have understood the nature of the research and why  
my son/daughter (circle one) has been selected to participate.

I have had the chance to ask questions and to have them answered to my satisfaction.

After reading through the Participant Information Sheet I do **NOT** agree that my child  
can participate in this research.

I understand that by signing this consent form, my child will **NOT** be participating in  
the research study.

I understand that I can request that my child be withdrawn from participation in the  
study at any time without penalty and that they can withdraw any data that they have  
submitted until 31 December 2013.

I understand that the school principal has made the assurance that the students' decision whether or not to participate in the study will have no impact on their learning, achievement or standing at school.

I understand that my child's personal information will remain confidential to the research team listed above and will not appear in any publications or reports that result from this research project.

I understand that my child will not be identified in any publications or reports that arise from this research.

I understand that all data (questionnaires and Consent Forms) will be kept locked up for the duration of the research project on the University of Auckland premises or in an approved long-term storage facility and the hardcopy data will be destroyed after six years. The digital data will be securely deleted from the university computer system.

I understand that the data from this research will be used in conference presentations, academic journal articles, a confidential report for The Foundation for Youth Development, a summary report of key findings for control schools and in the student researcher's PhD dissertation.

Name: \_\_\_\_\_ (please print)

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number  
2013/8912

## Appendix N



School of Learning, Development and Professional  
Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT CONSENT FORM

BOARD OF TRUSTEES/PRINCIPAL

**THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS**

**Title of Project:** Can it? Conducting an implementation evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand

**Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-Investigator) and Ms. Rachel Williams (University of Auckland PhD student)

I, ....., (print name) have read the Participation Information Sheet that I was given and have understood the nature of the research.

I have had a chance to ask questions and have them answered to my satisfaction.

I agree that the researcher, Rachel Williams, has permission to enter the school and interact with students in the Year .... class on ..... (insert date).

I understand that the researcher will be present in the school for approximately two hours on the ..... (insert date).

I understand that I can request that the school be withdrawn from participation in the study at any time without penalty.

I understand that the student's personal information will remain confidential to the research team listed above and will not appear in any publications or reports that result from this research project.

I understand that neither I, as principal, nor the school will be identified in any publications or reports that arise from this research.

I understand that all data (questionnaires and Consent Forms) will be kept locked up for the duration of the research project on the University of Auckland premises or in an approved long-term storage facility and will be destroyed after six years.

I understand that the data from this research will be used in conference presentations, academic journal articles, a confidential report for The Foundation for Youth Development and in the student researcher's PhD dissertation.

Name: \_\_\_\_\_

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number  
2013/8912

## Appendix O



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### Measure of General Knowledge

#### Part A. Personal Information

Please read the following questions and answer them by shading in the right box.

Example: Do you like ice cream?

- Yes  
 No

1. Are you a boy or a girl?

- Boy  
 Girl

2. What is your age?

- 10 years of age  
 11 years of age  
 12 years of age  
 13 years of age  
 14 years of age

3. What is your Year Level?

- Year 6  
 Year 7  
 Year 8

4. How long have you been at this school?

- 1 year  
 2 years  
 3 years  
 4 years  
 5 years  
 6 years  
 7 years  
 8 years  
 9 years

5. Which people group(s) do you belong to? (You can shade more than more box).

- Māori  
 Pasifika  
 Pakeha/New Zealand European  
 Asian  
 Other

6. Is Kiwi Can in your school?  
 Yes  
 No – If you answered “no” please move to Part B

7. Do you go to Kiwi Can classes each week?  
 Yes  
 No

**Part B. General Knowledge Quiz**

Please read each question below and write your answer in the space provided.

**Science**

	Question	Answer
1	What is the biggest planet in our solar system?	
2	What is the chemical symbol for the element oxygen?	
3	How many bones do sharks have in their bodies?	
4	True or false? Dogs are herbivores	
5	What is the name of the long appendage that hangs from an elephants face?	

	Question	Answer
6	The highest mountain on earth is?	
7	What is the name of the closest star to the earth?	
8	True or false? Frogs are cold blooded animals	
9	True or false? Yogurt is produced by bacterial fermentation of milk.	
10	True or false? The fastest land animal in the world is the zebra	

**Arts and Literature**

	Question	Answer
1	Who had an encounter with the three bears?	
2	What is the name of Winnie the Pooh’s donkey friend?	
3	Who was created by Gepetto the woodcarver?	
4	Who is the bear in The Jungle Book?	
5	What kind of creature was defeated by the Three Billy Goats Gruff?	

6	Who painted the Mona Lisa?	
7	Who leads a gang of outlaws in Sherwood Forest?	
8	Who stole Christmas in a Dr Seuss book?	
9	Who were all the King's horses and men unable to put together again?	
10	Who is the druid in the Asterix books?	

### Entertainment

	Question	Answer
1	Sheriff Woody Pride and Buzz Lightyear are major characters in which series of films?	
2	What is the name of Donald Duck's girlfriend?	
3	In which town do the Flintstones live?	
4	Leona Lewis and Alexandra Burke are former winners of which competition?	
5	What is the name of Shrek's wife?	

	Question	Answer
6	Wolverine is the leader of which band of superheroes?	
7	Who fell in love with a robot called EVE?	
8	What is the name of the vehicle in which Scooby Doo and his friends travel?	
9	Which stringed instrument is bigger than a violin but smaller than a cello?	
10	What was the name of the Kung Fu Panda in the film of that name?	

### Geography and Travel

	Question	Answer
1	At more than 6.5 million square miles, what is the largest country in the world?	
2	And which has the largest population, at more than 1.3 billion?	
3	What is the capital city of Spain?	
4	Mumbai, Chennai and Kolkata are major cities in which country?	
5	Which sea separates Europe from Africa?	

6	Which desert covers much of northern Africa?	
7	Which island country lies off China, Korea and Russia?	
8	Which ocean lies between Africa and Australia and south of Asia?	
9	In which mountain range is Mount Everest?	
10	Which river rises in Peru, enters the sea in Brazil and at no point is crossed by a bridge?	

## Nature

	Question	Answer
1	From what tree do acorns come?	
2	What colour are dandelions?	
3	How many ounces are there in a pound?	
4	What term in mathematics refers to the number obtained when two or more numbers are multiplied?	
5	How many sides does a hexagon have?	

## Appendix P



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### PARTICIPANT CONSENT FORM

#### TEACHER – CONTROL SCHOOL

#### THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

**Title of Project:** Can it? Conducting an outcome evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand

**Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-Investigator) and Ms. Rachel Williams (University of Auckland PhD student)

I, ....., (print name) have read the Participation Information Sheet that I was given and have understood the nature of the research.

I have had a chance to ask questions and have them answered to my satisfaction.

I agree to administer the questionnaire to the students in my Year \_\_\_\_ (fill in Year group) class during Week 2 or 3 of Term I and Week 6 or 7 of Term IV during the 2013 school year.

I understand that the researcher will provide me with a set of written instructions for administering the questionnaire and will visit me personally prior to the first date of administration.

I understand that I am responsible for handing out the Parent and Student Information Sheets, Parental Consent Forms and Student Assent Forms in Week 1 of Term 1 (2013).

I agree to collect all returned Parent Consent and Student Assent Forms and to store them in a safe and secure location prior to mailing them back to the research team in February 2013.

I understand that I can request that my students be withdrawn from participation in the study at any time without penalty and can withdraw data that they submit until 31 December 2013.

I understand that my principal has made the assurance that the students' decision whether or not to participate in the study will have no impact on their learning, achievement or standing at school.

I understand that the student's personal information will remain confidential to the research team listed above and will not appear in any publications or reports that result from this research project.

I understand that neither I, as the classroom teacher, nor the school will be identified in any publications or reports that arise from this research.

I understand that all hardcopy data (consent forms and questionnaires) will be kept locked in a cabinet in the office of the Principal Investigator at the University of Auckland or in an approved long-term storage facility for six years and will be destroyed after six years. The digital data will be permanently deleted from the university computer system.

I understand that the data from this research will be used in conference presentations, academic journal articles, a confidential report for The Foundation for Youth Development, a summary report of key findings for control schools and in the student researcher's PhD dissertation.

Name: \_\_\_\_\_ (please print)

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number  
2013/8912

## Appendix Q



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### TEACHER CONFIDENTIALITY AGREEMENT – CONTROL SCHOOLS

**Project Title:** Can it? An outcome evaluation of *Kiwi Can*, a school-based values and life skills programme in Aotearoa New Zealand

**Researcher:** Ms. Rachel Williams

I agree to store the Parental Consent Forms and Student Assent Forms that I receive in a secure location and understand that I am not permitted to discuss their presence with anyone other than the researchers.

I also agree not to view nor discuss any of the student's responses from the questionnaires that I administer and collect in February 2013 and November 2013.

<b>Name:</b>	
<b>Date:</b>	
<b>Signed:</b>	

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number 2013/8912

## Appendix R



School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### **PARTICIPANT CONSENT FORM** **BOARD OF TRUSTEES/PRINCIPAL**

**THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS**

**Title of Project:** Can it? Conducting an outcome evaluation of *Kiwi Can*, a school-based values and life-skills programme, in Aotearoa New Zealand

**Research Team:** Associate Professor Gavin Brown (Principal Investigator), Dr. Pat Bullen (Co-Investigator) and Ms. Rachel Williams (University of Auckland PhD student)

I, ....., (print name) have read the Participation Information Sheet that I was given and have understood the nature of the research.

I have had a chance to ask questions and have them answered to my satisfaction.

I agree that the researcher, Rachel Williams, has permission to enter the school and interact with teachers and students in our Year 6-8 classes on .....  
(insert date).

I understand that the researcher will be present in the school for approximately two hours on the ..... (insert date).

I understand that I can request that the school be withdrawn from participation in the study at any time without penalty and that the participants can withdraw data that they submit until 31 December 2013.

I understand that the student's personal information will remain confidential to the research team listed above and will not appear in any publications or reports that result from this research project.

I understand that neither I, as principal, nor the school will be identified in any publications or reports that arise from this research.

I understand that all data (questionnaires and Consent Forms) will be kept locked up for the duration of the research project on the University of Auckland premises or in an approved long-term storage facility and will be destroyed after six years. The digital data will be securely deleted from the university computer system.

I understand that the data from this research will be used in conference presentations, academic journal articles, a confidential report for The Foundation for Youth Development, a summary report of key findings for control schools and in the student researcher's PhD dissertation.

Yes, I wish to receive a summary report of key findings

Name: \_\_\_\_\_ (please print)

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number  
2013/8912

## Appendix S



THE UNIVERSITY OF AUCKLAND  
NEW ZEALAND

School of Learning, Development and Professional Practice  
Faculty of Education  
Epsom Campus  
Ph: 623 8899  
The University of Auckland  
Private Bag 92601  
Auckland, New Zealand

### Participation / Non-participation Assurance

Board of Trustees/Principal

**Project title:** Can it? An outcome evaluation of *Kiwi Can*, a school-based values and life skills programme, in Aotearoa New Zealand

**Name of Researcher:** Rachel Williams

I give assurance that the students' decision whether or not to participate in the study will have no impact on their learning, achievement or standing at this school.

Name \_\_\_\_\_(please print)

Signature \_\_\_\_\_ Date \_\_\_\_\_

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS  
ETHICS COMMITTEE ON 14 JANUARY 2013 for (3) years, Reference Number  
2013/8912

## Appendix T

### *Means, Standard Deviations, and Confidence Intervals for PYD Subscales at Baseline as a Function of School Condition, Sex, Year Level, and Ethnicity*

PYD Subscale	Category	<i>n</i>	<i>M</i>	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
Composite						
EC	Composite	1219	5.09	.86	5.04	5.13
PT	Composite	1219	4.59	.88	4.54	4.64
SRP	Composite	1219	5.31	.68	5.27	5.35
CS	Composite	1219	5.05	.79	5.01	5.09
CP	Composite	1219	4.82	.95	4.77	4.89
VD	Composite	1219	4.84	.84	4.79	4.88
PV	Composite	1219	5.00	.80	4.95	5.05
Condition						
EC	Experienced	299	4.98	.87	4.88	5.07
	New	464	5.09	.90	5.01	5.17
	Control	456	5.15	.80	5.08	5.23
PT	Experienced	299	4.44	.90	4.34	4.54
	New	464	4.66	.88	4.58	4.74
	Control	456	4.61	.86	4.53	4.70
SRP	Experienced	299	5.20	.68	5.12	5.27
	New	464	5.35	.67	5.29	5.41
	Control	456	5.34	.67	5.28	5.41
CS	Experienced	299	4.97	.78	4.88	5.06
	New	464	5.03	.82	4.96	5.11
	Control	456	5.12	.77	5.05	5.12
CP	Experienced	299	4.70	.96	4.59	4.81
	New	464	4.80	1.01	4.71	4.88
	Control	456	4.92	.87	4.83	5.01
VD	Experienced	299	4.79	.82	4.69	4.88
	New	464	4.82	.87	4.75	4.90
	Control	456	4.88	.82	4.80	4.96
PV	Experienced	299	4.88	.84	4.79	4.97
	New	464	5.06	.79	4.99	5.13
	Control	456	5.02	.78	4.95	5.09
Sex						
EC	boy	606	4.88	.96	4.81	4.95
	girl	613	5.29	.70	5.22	5.36
PT	boy	606	4.41	.88	4.34	4.48
	girl	613	4.77	.84	4.70	4.84
SRP	boy	606	5.16	.75	5.12	5.21
	girl	613	5.46	.55	5.40	5.51
CS	boy	606	4.97	.82	4.90	5.03
	girl	613	5.13	.75	5.07	5.20
CP	boy	606	4.73	.99	4.65	4.80
	girl	613	4.91	.90	4.84	4.99
VD	boy	606	4.69	.91	4.62	4.75
	girl	613	4.98	.73	4.92	5.05
PV	boy	606	4.84	.86	4.78	4.91
	girl	613	5.16	.71	5.09	5.22
Year Level						
EC	Year 6	371	5.14	.88	5.06	5.23
	Year 7	573	5.14	.85	5.07	5.21
	Year 8	273	4.89	.84	4.79	4.99
PT	Year 6	371	4.62	.88	4.54	4.71
	Year 7	573	4.66	.88	4.59	4.74

PYD Subscale	Category	<i>n</i>	<i>M</i>	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
SRP	Year 8	273	4.39	.87	4.29	4.49
	Year 6	371	5.34	.70	5.27	5.41
	Year 7	573	5.36	.69	5.30	5.41
CS	Year 8	273	5.16	.70	5.08	5.24
	Year 6	371	5.23	.68	5.15	5.31
	Year 7	573	5.06	.74	4.50	5.13
CP	Year 8	273	4.78	.77	4.69	4.87
	Year 6	371	4.88	.90	4.79	4.98
	Year 7	573	4.82	.98	4.74	4.89
VD	Year 8	273	4.74	.94	4.62	4.85
	Year 6	371	4.97	.82	4.88	5.05
	Year 7	573	4.86	.84	4.79	4.93
PT	Year 8	273	4.61	.82	4.51	4.71
	Year 6	371	5.10	.78	5.02	5.18
	Year 7	573	5.05	.77	4.98	5.11
Ethnicity	Year 8	273	4.77	.86	4.68	4.87
	EC					
	Māori	454	4.95	.96	4.87	5.03
PT	Pasifika	117	5.32	.79	5.16	5.47
	Pakeha	515	5.13	.74	5.06	5.20
	Asian	46	5.13	.92	4.88	5.37
	Other	82	5.21	1.01	5.03	5.40
	Māori	454	4.47	.94	4.39	4.55
SRP	Pasifika	117	4.67	.84	4.51	4.87
	Pakeha	515	4.66	.81	4.58	4.73
	Asian	46	4.67	.85	4.42	4.93
	Other	82	4.69	1.01	4.50	4.88
	Māori	454	5.17	.75	5.11	5.23
CS	Pasifika	117	5.43	.57	5.31	5.55
	Pakeha	515	5.40	.54	5.31	5.46
	Asian	46	5.40	.64	5.20	5.59
	Other	82	5.28	.97	5.14	5.43
	Māori	454	4.96	.88	4.89	5.03
CP	Pasifika	117	5.26	.64	5.12	5.40
	Pakeha	515	5.06	.72	4.99	5.13
	Asian	46	5.21	.59	4.98	5.44
	Other	82	5.11	.99	4.93	5.28
	Māori	454	4.75	1.04	4.66	4.83
VD	Pasifika	117	4.71	1.04	4.54	4.88
	Pakeha	515	4.88	.85	4.80	4.96
	Asian	46	4.98	.65	4.71	5.25
	Other	82	4.91	.98	4.71	5.11
	Māori	454	4.78	.87	4.71	4.86
PV	Pasifika	117	5.32	.62	5.18	5.47
	Pakeha	515	4.72	.79	4.65	4.79
	Asian	46	5.29	.50	5.06	5.53
	Other	82	4.92	1.05	4.74	5.09
	Māori	454	4.87	.89	4.80	4.95
PV	Pasifika	117	5.23	.70	5.08	5.37
	Pakeha	515	5.04	.68	4.97	5.11
	Asian	46	5.16	1.00	4.93	5.39
	Other	82	5.09	.80	4.92	5.27

*Note.* EC = empathic concern; PT = perspective taking; SRP = social rules/politeness; CS = connection to school; CP = connection to peers; VD = values diversity; PV = personal values; Mean scores based on balanced 6-point Likert scale (1 = *strongly disagree* to 6 = *strongly agree*).

## Appendix U

### *Means, Standard Deviations and Confidence Intervals for CC Subscales at Baseline as a Function of School Condition, Sex, Year Level, and Ethnicity*

CC Subscale	Category	<i>n</i>	<i>M</i>	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
Composite						
PBCSS	Composite	1219	4.73	.80	4.68	4.77
NB	Composite	1219	3.35	1.11	3.29	3.41
SS	Composite	1219	4.94	.90	4.89	4.99
ES	Composite	1219	4.97	1.07	4.91	5.03
Condition						
PBCSS	Experienced	299	4.62	.82	4.53	4.72
	New	464	4.75	.82	4.68	4.82
	Control	456	4.76	.80	4.69	4.83
NB	Experienced	299	3.91	1.07	3.79	4.03
	New	464	3.64	1.10	3.54	3.73
	Control	456	3.49	1.10	3.39	3.59
SS	Experienced	299	4.81	.92	4.71	4.92
	New	464	4.88	.95	4.80	4.96
	Control	456	5.04	.84	4.95	5.12
ES	Experienced	299	4.82	1.19	4.69	4.94
	New	464	5.01	.99	4.92	5.11
	Control	456	4.96	1.07	4.86	5.06
Sex						
PBCSS	boy	606	4.63	.86	4.56	4.69
	girl	613	4.82	.75	4.75	4.88
NB	boy	606	3.70	1.08	3.61	3.79
	girl	613	3.59	1.21	3.51	3.68
SS	boy	606	4.87	.95	4.79	4.94
	girl	613	4.98	.86	4.90	5.05
ES	boy	606	4.82	1.10	4.74	4.91
	girl	613	5.06	1.03	4.98	5.15
Year Level						
PBCSS	Year 6	371	4.82	.79	4.74	4.90
	Year 7	573	4.76	.82	4.69	4.82
	Year 8	273	4.52	.80	4.43	4.62
NB	Year 6	371	3.75	1.15	3.63	3.86
	Year 7	573	3.53	1.09	3.44	3.62
	Year 8	273	3.76	1.04	3.63	3.89
SS	Year 6	371	5.11	.83	5.01	5.20
	Year 7	573	4.92	.94	4.84	4.99
	Year 8	273	4.68	.88	4.57	4.79
ES	Year 6	371	5.17	.99	5.06	5.28
	Year 7	573	5.01	1.00	4.92	5.09
	Year 8	273	4.50	1.19	4.36	4.65
Ethnicity						
PBCSS	Māori	454	4.62	.87	4.55	4.69
	Pasifika	117	4.99	.67	4.85	5.14
	Pakeha	515	4.75	.75	4.68	4.82
	Asian	46	4.61	.81	4.38	4.84
	Other	82	4.81	.93	4.64	4.99
NB	Māori	454	3.92	1.10	3.82	4.02
	Pasifika	117	3.63	1.13	3.43	3.82
	Pakeha	515	3.46	1.04	3.37	3.56
	Asian	46	3.51	1.10	3.20	3.83

CC Subscale	Category	<i>n</i>	<i>M</i>	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
SS	Other	82	3.41	1.10	3.18	3.65
	Māori	454	4.83	.97	4.75	4.91
	Pasifika	117	5.06	.82	4.90	5.23
	Pakeha	515	4.99	.81	4.91	5.06
	Asian	46	4.96	.88	4.70	5.22
ES	Other	82	4.86	1.09	4.66	5.05
	Māori	454	4.85	1.14	4.75	4.94
	Pasifika	117	5.39	.86	5.20	5.58
	Pakeha	515	4.89	1.02	4.80	4.98
	Asian	46	5.30	.73	4.99	5.60
	Other	82	5.04	1.07	4.81	5.27

*Note.* PBCSS = positive behaviour and classroom supportiveness; NB = negative behaviour; SS = safety at school; ES = enjoyment at school. Mean scores based on balanced 6-point Likert scale (1 = *strongly disagree* to 6 = *strongly agree*).

## Appendix V

### *Mean Change Scores, Standard Deviations, and Confidence Intervals for PYD*

#### *Subscales*

Dependent Variable	Category	<i>n</i>	$\Delta M$	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
<b>Composite</b>						
EC	Composite	1219	-.20	.88	-.25	-.15
PT	Composite	1219	-.07	.89	-.12	-.02
SRP	Composite	1219	-.11	.68	-.14	-.06
CS	Composite	1219	-.24	.85	-.29	-.19
CP	Composite	1219	-.19	1.00	-.25	-.13
VD	Composite	1219	-.08	.85	-.13	-.03
PV	Composite	1219	-.10	.81	-.14	-.05
<b>Condition</b>						
EC	Experienced	299	-.34	.96	-.44	-.24
	New	464	-.19	.87	-.28	-.12
	Control	456	-.12	.83	-.21	-.05
PT	Experienced	299	-.03	.97	-.13	.07
	New	464	-.09	.82	-.17	-.01
	Control	456	-.08	.90	-.16	.00
SRP	Experienced	299	-.16	.74	-.24	-.08
	New	464	-.07	.61	-.13	-.01
	Control	456	-.09	.71	-.16	-.03
CS	Experienced	299	-.33	.92	-.43	-.23
	New	464	-.16	.79	-.24	-.08
	Control	456	-.27	.85	-.35	-.19
CP	Experienced	299	-.21	1.01	-.32	-.10
	New	464	-.15	1.00	-.24	-.06
	Control	456	-.27	.99	-.31	-.13
VD	Experienced	299	-.15	.94	-.25	-.06
	New	464	-.08	.82	-.15	.00
	Control	456	-.04	.81	-.11	-.13
PV	Experienced	299	-.15	.91	-.24	-.06
	New	464	-.12	.74	-.20	-.05
	Control	456	-.05	.81	-.12	.03
<b>Sex</b>						
EC	boy	606	-.19	.94	-.26	-.12
	girl	613	-.22	.81	-.28	-.14
PT	boy	606	-.04	.91	-.12	.03
	girl	613	-.10	.87	-.17	-.03
SRP	boy	606	-.06	.75	-.12	-.01
	girl	613	-.14	.61	-.19	-.08
CS	boy	606	-.22	.88	-.29	-.15
	girl	613	-.26	.81	-.33	-.19
CP	boy	606	-.15	1.01	-.23	-.07
	girl	613	-.24	.98	-.32	-.16
VD	boy	606	-.03	.90	-.10	.04
	girl	613	-.13	.80	-.19	-.06
PV	boy	606	-.07	.87	-.13	.00
	girl	613	-.13	.74	-.20	-.07
<b>Year Level</b>						
EC	Year 6	371	-.24	.94	-.33	-.15
	Year 7	573	-.22	.85	-.29	-.15
	Year 8	273	-.12	.86	-.22	-.01
PT	Year 6	371	-.01	.94	-.11	.08

Dependent Variable	Category	n	$\Delta M$	SD	95% CI	
					LL	UL
SRP	Year 7	573	-.12	.86	-.20	-.05
	Year 8	273	-.04	.89	-.15	.06
	Year 6	371	-.08	.71	-.15	.01
CS	Year 7	573	-.10	.64	-.15	-.04
	Year 8	273	-.13	.74	-.22	-.05
	Year 6	371	-.22	.87	-.31	-.14
CP	Year 7	573	-.23	.80	-.30	-.16
	Year 8	273	-.13	.91	-.40	-.19
	Year 6	371	-.19	.93	-.29	-.09
VD	Year 7	573	-.17	1.02	-.25	-.09
	Year 8	273	-.24	1.04	-.36	-.12
	Year 6	371	-.09	.92	-.17	.00
PT	Year 7	573	-.10	.80	-.17	-.03
	Year 8	273	-.04	.86	-.14	.07
	Year 6	371	-.09	.85	-.17	-.01
Ethnicity	Year 7	573	-.12	.75	-.19	-.06
	Year 8	273	-.08	.87	-.17	.03
	Year 6	371	-.09	.85	-.17	-.01
EC	Māori	454	-.17	.94	-.26	-.09
	Pasifika	117	-.32	.90	-.47	-.15
	Pakeha	515	-.22	.81	-.29	-.14
	Asian	46	-.13	.78	-.38	.13
	Other	82	-.12	.97	-.31	.07
PT	Māori	454	-.03	.92	-.11	.05
	Pasifika	117	-.12	.86	-.28	.04
	Pakeha	515	-.13	.85	-.21	-.05
	Asian	46	.09	.77	-.17	.34
	Other	82	.04	1.05	-.15	.24
SRP	Māori	454	-.11	.76	-.17	-.05
	Pasifika	117	-.15	.59	-.27	-.02
	Pakeha	515	-.13	.60	-.19	-.07
	Asian	46	.02	.51	-.18	.22
	Other	82	.11	.90	-.04	.26
CS	Māori	454	-.21	.93	-.29	-.13
	Pasifika	117	-.38	.75	-.54	-.23
	Pakeha	515	-.26	.78	-.34	-.19
	Asian	46	-.13	.63	-.38	.11
	Other	82	-.13	.98	-.31	.05
CP	Māori	454	-.12	1.08	-.21	-.03
	Pasifika	117	-.07	.91	-.25	.11
	Pakeha	515	-.30	.94	-.39	-.22
	Asian	46	-.22	.88	-.50	.07
	Other	82	-.05	1.05	-.26	.17
VD	Māori	454	-.09	.93	-.17	-.11
	Pasifika	117	-.18	.70	-.33	-.03
	Pakeha	515	-.08	.82	-.15	-.01
	Asian	46	.07	.54	-.30	.19
	Other	82	-.09	.92	-.11	.26
PV	Māori	454	-.11	.88	-.19	-.04
	Pasifika	117	-.16	.78	-.31	-.01
	Pakeha	515	-.13	.74	-.20	-.06
	Asian	46	.02	.95	-.21	.26
	Other	82	.11	.81	-.06	.29

Note. EC = empathic concern; PT = perspective taking; SRP = social rules/politeness; CS = connection to school; CP = connection to peers; VD = values diversity; PV = personal values. Mean scores based on balanced 6-point Likert scale (1=*strongly disagree* to 6 = *strongly agree*).

## Appendix W

### *Mean Change Scores, Standard Deviations, and Confidence Intervals for CC Subscales*

CC Subscale	Category	<i>n</i>	$\Delta M$	<i>SD</i>	95% CI <i>LL</i>	<i>UL</i>
Composite						
PBCSS	Composite	1219	-.25	.86	-.30	-.20
NB	Composite	1219	.35	1.15	.28	.41
SS	Composite	1219	-.11	1.04	-.18	-.05
ES	Composite	1219	-.37	1.12	-.43	-.30
Condition						
PBCSS	Experienced	299	-.24	.87	-.34	-.15
	New	464	-.24	.86	-.32	-.16
	Control	456	-.26	.87	-.34	-.18
NB	Experienced	299	.20	1.23	.08	.34
	New	464	.40	1.10	.30	.51
	Control	456	.38	1.14	.27	.49
SS	Experienced	299	-.14	1.15	-.26	-.12
	New	464	-.03	1.14	-.13	.06
	Control	456	-.17	1.01	-.27	-.08
ES	Experienced	299	-.50	1.36	-.62	-.37
	New	464	-.30	1.04	-.40	-.20
	Control	456	-.36	1.03	-.47	-.26
Sex						
PBCSS	boy	606	-.24	.91	-.31	-.17
	girl	613	-.26	.82	-.33	-.19
NB	boy	606	.35	1.16	.26	.44
	girl	613	.34	1.13	.25	.43
SS	boy	606	-.16	1.11	-.24	-.07
	girl	613	-.07	.98	-.15	.02
ES	boy	606	-.40	1.20	-.48	-.30
	girl	613	-.35	1.04	-.44	-.26
Year Level						
PBCSS	Year 6	371	-.18	.86	-.27	-.09
	Year 7	573	-.30	.85	-.37	-.23
	Year 8	273	-.24	.91	-.34	-.14
NB	Year 6	371	.32	1.24	.20	.43
	Year 7	573	.40	1.13	.31	.50
	Year 8	273	.27	1.07	.14	.41
SS	Year 6	371	-.10	.98	-.20	.01
	Year 7	573	-.10	1.04	-.18	-.01
	Year 8	273	-.15	1.13	-.27	-.02
ES	Year 6	371	-.35	1.14	-.46	-.23
	Year 7	573	-.41	1.06	-.50	-.31
	Year 8	273	-.34	1.23	-.47	-.21
Ethnicity						
PBCSS	Māori	454	-.15	.91	-.23	-.08
	Pasifika	117	-.27	.71	-.42	-.11
	Pakeha	515	-.36	.83	-.43	-.28
	Asian	46	-.05	.76	-.30	.20
	Other	82	-.17	.99	-.36	.01
NB	Māori	454	.20	1.17	.09	.30
	Pasifika	117	.43	1.00	.22	.64
	Pakeha	515	.45	1.11	.34	.54
	Asian	46	.44	1.10	.11	.77
	Other	82	.40	1.40	.15	.65

CC Subscale	Category	<i>n</i>	$\Delta M$	<i>SD</i>	95% CI	
					<i>LL</i>	<i>UL</i>
SS	Māori	454	-.08	1.10	-.17	.02
	Pasifika	117	-.16	1.01	-.35	.03
	Pakeha	515	-.18	.96	-.27	-.09
	Asian	46	.11	.91	-.19	.41
	Other	82	.12	1.28	-.11	.35
ES	Māori	454	-.32	1.20	-.42	-.21
	Pasifika	117	-.41	1.03	-.61	-.20
	Pakeha	515	-.47	1.07	-.57	-.37
	Asian	46	-.19	.77	-.51	.14
	Other	82	-.11	1.22	-.35	.13

*Note.* PBCSS = positive behaviour and classroom supportiveness; NB = negative behaviour; SS = safety at school; ES = enjoyment at school. Mean scores based on balanced 6-point Likert scale (1 = *strongly disagree* to 6 = *strongly agree*).

## References

- Aber, J. L., Bennett, N. G., Conley, D.C., & Li, J. (1997). The effects of poverty on child health and development. *Annual Review of Public Health, 18*(1), 463-483. doi:10.1146/annurev.publhealth.18.1.463
- Abu-Hilal, M. (1992). A structural model for the direct and indirect effects of sex, math anxiety on attitudes and achievement in mathematics. *Journal of Faculty of Education (Almansoura University), 20*, 35-53. Retrieved from <https://scholar.google.co.nz/citations?user=Jkh6N2MAAAAJ&hl=en&oi=sra>
- Ackerman, B. P., Kogos, J., Youngstrom, E., Schoff, K., & Izard, C. (1999). Family instability and the problem behaviors of children from economically disadvantaged families. *Developmental Psychology, 35*(1), 258-268. doi:10.1037/0012-1649.35.1.258
- Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family, 67*(4), 1012-1028. doi:10.1111/j.1741-3737.2005.00191.x
- Adolescent Health Research Group. (2013). *The health and well-being of New Zealand secondary school students in 2012: Youth '12 prevalence tables*. Auckland, New Zealand: The University of Auckland. Retrieved from <https://cdn.auckland.ac.nz/assets/fmhs/faculty/ahrg/docs/2012prevalence-tables-report.pdf>
- Ahmed, S. (2010). *Promoting positive development in school children: Perspectives of the Kiwi Can programme in New Zealand*. (Unpublished masters' thesis). University of Auckland, Auckland, New Zealand.
- Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behaviour and Human Decision Processes, 50*(2), 179-211. doi:10.1016/0749-5978(91)90020-T
- Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology and Health, 26*(9), 1113-1127. doi:10.1080/08870446.2011.613995
- Akos, P. (2002). Student perceptions of the transition from elementary to middle school. *Professional School Counseling, 5*(5), 339-345. Retrieved from <http://www.biomedsearch.com/article/Student-perceptions-transition-from-elementary/88579043.html>
- Aldenderfer, M. S., & Blashfield, R. K. (1984). Cluster analysis. In M. S. Lewis-Beck (Ed.), *Quantitative applications in the social sciences* (pp. 1-88). Newbury Park, CA: Sage.
- Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1996). Children in motion: School transfers and elementary school performance. *Journal of Educational Research, 90*(1), 3-12. doi:10.1080/00220671.1996.9944438
- Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences, 32*(3-4), 249-311. doi:10.1017/S0140525X09990951
- Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation, 31*(3), 363-381. doi:10.1177/1098214010371972
- Atkinson, J., Salmond, C., & Crampton, P. (2014). *NZDep2013 index of deprivation*. Wellington, New Zealand: University of Otago.
- Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research, 24*(2), 222-228. doi:10.2307/3151512

- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Monterey, CA: Brooks Cole.
- Barlow, A. (1990). *Curriculum issues in health education in primary and secondary schools in New Zealand*. (Unpublished master's thesis). University of Auckland, Auckland, New Zealand.
- Barnard, H., & Turner, C. (2011). *Poverty and ethnicity: A review of evidence*. London, England: Joseph Rowntree Foundation. Retrieved from <http://www.jrf.org.uk/sites/files/jrf/poverty-ethnicity-evidence-summary.pdf>
- Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behaviour game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behaviour Analysis*, 2(2), 119-124. doi: 10.1901/jaba.1969.2-119
- Baulch, B., Nguyen, H. T. M., Phuong, P. T. T., & Pham, H. T. (2010). *Ethnic minority poverty in Vietnam: Working paper No. 169*. Manchester, England: Chronic Poverty Research Centre. Retrieved from [http://www.chronicpoverty.org/uploads/publication\\_files/WP169%20Baulch%20etal.pdf](http://www.chronicpoverty.org/uploads/publication_files/WP169%20Baulch%20etal.pdf)
- Belle, D. E. (1983). The impact of poverty on networks and supports. *Marriage & Family Review*, 5(4), 89-103. doi:10.1300/J002v05n04\_06
- Benson, P. L. (1990). *The troubled journey: A portrait of 6th–12th grade youth*. Minneapolis, MN: Search Institute.
- Benson, P. L. (1997). *All kids are our kids: What communities must do to raise caring and responsible children and adolescents*. San Francisco, CA: Jossey-Bass.
- Benson, P. L. (2006). *All kids are our kids: What communities must do to raise caring and responsible children and adolescents* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Benson, P. L., & Leffert, N. (2001). Childhood and adolescence: Developmental assets. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 1690-1697). Irvine, CA: Elsevier.
- Benson, P. L., Leffert, N., Scales, P. C., & Blyth, D. A. (1998). Beyond the village rhetoric: Creating healthy communities for children and adolescents. *Applied Developmental Science*, 2(3), 138-159. doi:10.1207/s1532480xads0203\_3
- Benson, P. L., & Scales, P. C. (2009). The definition and preliminary measurement of thriving in adolescence. *The Journal of Positive Psychology*, 4(1), 85-104. doi:10.1080/17439760802399240
- Benson, P. L., Scales, P. C., Hamilton, S. F., Sesma Jr., A., Hong, K. L., & Roehlkepartain, E. C. (2006). Positive youth development so far: Core hypothesis and their implications for policy and practice. *Search Institute: Insights and Evidence*, 3(1), 1-13. Retrieved from <http://www.isbe.state.il.us/learningsupports/climate/pdfs/positive-youth-dev.pdf>
- Benson, P. L., Scales, P. C., Leffert, N., Roehlkepartain, E. C. (1999). *A fragile foundation: The state of development assets among American youth*. Minneapolis, MN: Search Institute.
- Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles, CA: BMDP Statistical Software.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246. doi:10.1037/0033-2909.107.2.238
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

- Bentler, P. M. (2005). *EQS6 structural equations program manual*. Encino, CA: Multivariate Software.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588-606. doi:10.1037/0033-2909.88.3.588
- Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual Review of Psychology*, 47(1), 541-570. doi:10.1146/annurev.psych.47.1.563
- Bergstrom, E. (1988). Health teams in school: A step towards primary prevention. *Pediatrician*, 15(3), 137-142. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3217292>
- Bevans, K. B., Bradshaw, C. P., Miech, R., & Leaf, P. J. (2007). Staff- and school-level predictors of school organizational health: A multilevel analysis. *Journal of School Health*, 77(6), 294-302. doi:10.1111/j.1746-1561.2007.00210.x
- Birnbaum, A. S., Lytle, L. A., Hannan, P. J., Murray, D. M., Perry, C. L., & Forester, J. L. (2003). School functioning and violent behavior among young adolescents: A contextual analysis. *Health Education Research*, 18(3), 389-403. doi:10.1093/her/cyf036
- Blalock, H. M. (1979). *Social statistics* (2nd ed.). Tokyo, Japan: McGraw-Hill Kogakusha.
- Blane, D., Berney, L., Davey-Smith, G., Gunnell, D., & Holland, P. (1999). Reconstructing the life course: A 60-year follow-up study based on the Boyd Orr cohort. *Public Health*, 113(3), 117-124. doi:10.1038/sj.ph.1900551
- Boheim, R., & Taylor, M. P. (2002). Tied down or room to move? Investigating the relationships between housing tenure, employment status, and residential mobility in Britain. *Scottish Journal of Political Economy*, 49(4), 369-392. doi:10.1111/1467-9485.00237
- Bollen, K. A. (1989). *Structural equation modelling with latent variables*. New York, NY: Wiley.
- Borg, W. R., & Gall, M. D. (1979). *Educational research: An introduction* (6th ed.). New York, NY: Longman.
- Boston, J. (2013, October). *Child poverty in New Zealand: Why it matters and how it can be reduced*. Paper presented at Children in Crisis Conference, University of Waikato, Hamilton, New Zealand. Retrieved from [http://igps.victoria.ac.nz/documents/Child\\_poverty\\_in\\_New\\_Zealand\\_Why\\_it\\_mattersOct13.pdf](http://igps.victoria.ac.nz/documents/Child_poverty_in_New_Zealand_Why_it_mattersOct13.pdf)
- Bowers, E. P., Li, Y., Kiely, M. K., Brittan, A., Lerner, J. V., & Lerner, R. M. (2010). The five Cs model of positive youth development: A longitudinal analysis of confirmatory factor structure and measurement invariance. *Journal of Youth and Adolescence*, 39(7), 720-735. doi:10.1007/s10964-010-9530-9
- Bowers, E. P., von Eye, A., Lerner, J. V., Arbeit, M. R., Weiner, M. B., Chase, P., & Agans, J.P. (2011). The role of ecological assets in positive and problematic developmental trajectories. *Journal of Adolescence*, 34, 1151-1165. doi:10.1016/j.adolescence.2011.07.007
- Bowling, A. (2005). *Measuring health: A review of quality of life measurement scales* (3rd ed.). Berkshire, England: Open University Press.
- Bradshaw, C. P., Sawyer, A. L., & O'Brennan, L. M. (2009). A social disorganization perspective on bullying-related attitudes and behaviours: The influence of school context. *American Journal of Community Psychology*, 43(3), 204-220. doi:10.1007/s10464-009-9240-1

- Brazil, K. (1999). A framework for developing evaluation capacity in health care settings. *International Journal of Health Care Quality Assurance incorporating Leadership in Health Services*, 12(1), vi-xi. Retrieved from <http://www.emeraldinsight.com/doi/pdfplus/10.1108/13660759910249693>
- Breivik, E., & Olsson, U. H. (2001). Adding variables to improve fit: The effect of model size on fit assessment in LISREL. In R. Cudeck, R. G., Jöreskog, S. H. C. du Toit, & D. Sörbom (Eds.), *Structural equation modeling present and future: A festschrift in honor of Karl Jöreskog* (pp. 169-194). Chicago, IL: Scientific Software.
- Bristol, N. (2012). Preventing disease: History of disease prevention in the U.S. *CQ Researcher*, 22(1), 10-15. Retrieved from <http://photo.pds.org:5012/cqresearcher/getpdf.php?id=cqresrre2012010600>
- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In R. M. Lerner, *Handbook of child psychology* (Vol. 1, pp.793-828). Hoboken, NJ: John Wiley & Sons.
- Brown, G. T. L. (2004). Measuring attitude with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports*, 94(3), 1015-1024. doi:10.2466/pr0.94.3.1015-1024
- Brown, D., Benzeval, M., Gayle, V., Macintyre, S., O'Reilly, D., & Leyland, A. H. (2012). Childhood residential mobility and health in late adolescence and adulthood: Findings from the West Scotland Twenty-07 study. *Journal of Epidemiology and Community Health*, 66(10), 942-950. doi:10.1136/jech-2001-200316.
- Brown, G., Bhrolchain, M., & Harris, T. (1975). Social class and psychiatric disturbance among women in an urban population. *Sociology*, 9(2), 225-254. doi:10.1177/003803857500900203
- Brown, M. B., & Forsythe, A. B. (1974). The small sample behaviour of some statistics which test the equality of several means. *Technometrics*, 16(1), 129-132. doi:10.1080/00401706.1974.10489158
- Browne, M. W., & Cudeck, R. (1989). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2), 230-258. doi:10.1177/0049124192021002005
- Bruno, J. E., & Isken, J. (1996). Inter- and intra-school site student transiency: Practical and theoretical implications for instructional continuity at inner city schools. *Journal of Research and Development in Education*, 29(4), 239-248. Retrieved from <http://eric.ed.gov/?id=EJ535022>
- Burnham, K. P., & Anderson, D. R. (2004). Multi-model inference: Understanding AIC and BIC in model selection. *Sociological Methods Research*, 33(2), 261-304. doi:10.1177/0049124104268644
- Burns, R. B. (2000). *Introduction to research methods* (4th ed.). London, England: Pearson Education.
- Buston, K., Wight, D., Hart G., & Scott, S. (2002). Implementation of a teacher-delivered sex education programme: Obstacles and facilitating factors. *Health Education Research*, 17(1), 59-72. doi:10.1093/her/17.1.59
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.

- Campbell, D. T. (1970). Considering the case against experimental evaluations of social innovations. *Administrative Science Quarterly*, 15(1), 110-113. doi:10.2307/2391194
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally & Company.
- Carlo, G., Fabes, R. A., Laible, D., & Kupanoff, K. (1999). Early adolescence and prosocial behavior II: The role of social and contextual influences. *Journal of Early Adolescence*, 19(2), 133-147. doi:10.1177/0272431699019002001
- Catalano, R. F., Hawkins, J. D., Berglund, L. M., Ryan, J. A. M., & Lonczak, H. S. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Annals of the American Academy of Political and Social Science*, 591(1), 98-124. doi:10.1177/0002716203260102
- Catterall, J. S. (1998). Risk and resilience in student transitions to high school. *American Journal of Education*, 106(2), 302-334. doi:10.1086/444184
- Chang, L. (1994). A psychometric evaluation of four-point and six-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, 18(3), 205-215. doi:10.1177/014662169401800302
- Chen, H-T. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H-T. (1994). Current trends and future directions in program evaluation. *Evaluation Practice*, 15(3), 229-238. doi:10.1016/0886-1633(94)90016-7
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. doi:10.1207/S15328007SEM0902\_5
- Child Poverty Action Group [CPAG]. (2014a). *Our children, our choice: Priorities for policy*. Wellington, New Zealand: CPAG. Retrieved from <http://www.cpag.org.nz/assets/Publications/1-0%20Our%20Children%20Our%20Choice%20COMPLETE.pdf>
- Child Poverty Action Group [CPAG] (2014b). *The revolving door: Student mobility in Auckland schools*. Wellington, New Zealand: CPAG. Retrieved from <http://www.cpag.org.nz/assets/Backgrounders/140523%20CPAG%20Transience%20Report%20May2014%20FINAL.pdf>
- Child Poverty Monitor. (2014). *Measuring child poverty: Introduction*. Wellington, New Zealand: Child Poverty Monitor. Retrieved from [http://www.nzchildren.co.nz/document\\_downloads/B.%202014%20CPM%20Child%20Poverty%20and%20Living%20Standards.pdf](http://www.nzchildren.co.nz/document_downloads/B.%202014%20CPM%20Child%20Poverty%20and%20Living%20Standards.pdf)
- Children's Commission. (2012). *Solutions to child poverty in New Zealand: Evidence for action*. Wellington, New Zealand: Children's Commission of New Zealand. Retrieved from <http://www.occ.org.nz>
- Cholevas, N. K., & Loucaides, C. A. (2012). Factors that facilitate and barriers towards the implementation of health educational programmes in primary education schools of the prefecture of Achaia, Greece. *Health Education Journal*, 71(3), 365-375. doi:10.1177/0017896911406964
- Chou, C-P., Bentler, P. M., & Satorra, A. (1991). Scale test statistics and robust standard errors for nonnormal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology*, 44(2), 347-357. doi:10.1111/j.2044-8317.1991.tb00966.x

- Clark, T. C., Fleming, T., Bullen, P., Denny, S., Crengle, S., Dyson, B.,...Utter, J. (2013). *Youth '12 overview: The health and wellbeing of New Zealand secondary school students in 2012*. Auckland, New Zealand: The University of Auckland.
- Clark, T. C., Robinson, E., Crengle, S., Grant, S., Galbreath, R. A., & Sykora, J. (2009). *Youth '07: The health and wellbeing of secondary school students in New Zealand: Young people and violence*. Auckland, New Zealand: The University of Auckland.
- Clark, T. C., Robinson, E., Crengle, S., Herd, R., Grant, S., & Denny, S. (2008). *Te Ara Whakapiki Taitamariki. Youth '07: The health and wellbeing survey of secondary school students in New Zealand - Results for Māori young people*. Auckland, New Zealand: The University of Auckland.
- Cohen, L. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). New Jersey, NJ: Lawrence Erlbaum.
- Cohen, L. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. doi:10.1037/0033-2909.112.1.155
- Cohen, L., & Holliday, M. (1979). *Statistics for education and physical education*. London, England: Harper & Row.
- Cohen, L., & Holliday, M. (1982). *Statistics for social scientists*. London, England: Harper & Row.
- Cohen, L., & Holliday, M. (1996). *Practical statistics for students*. London, England: Paul Chapman Publishing.
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). New York, NY: RoutledgeFalmer.
- Collins, S. (2000, November 25). Clawbacks temper rent relief. *The New Zealand Herald*. Retrieved from [http://www.nzherald.co.nz/nz/news/article.cfm?c\\_id=1&objectid=161893](http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=161893)
- Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46(4), 648-659. doi:10.1037/0022-006X.46.4.648
- Conger, R. D., Ge, X., Elder, G. H., Lorenz, F. O., & Simons, R. L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, 65(2), 541-561. doi:10.2307/1131401
- Cook, T. D. (2000). The false choice between theory-based evaluation and experimentation. *New Directions for Evaluation*, 2000(87), 27-34. doi:10.1002/ev.1179
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design analysis issues for field settings*. Boston, MA: Houghton Mifflin.
- Cool Bananas. (2012). *Cool Bananas programme homepage*. Retrieved from <http://www.coolbananas.org.nz/programmes>
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5-23. doi:10.1002/ev.1114
- Craig, E., Jackson, C., & Han, D. (2007). *Monitoring the Health of New Zealand Children and Young People: Indicator Handbook*. Auckland, New Zealand: Paediatric Society of New Zealand and New Zealand Child and Youth Epidemiology Service. Retrieved from <http://www.paediatrics.org.nz/files/Indicator%20Handbook%20Version%2008.3.pdf>

- Craig, E., Reddington, A., Wicken, A., Oben, G., & Simpson, J. (2013). *Child poverty monitor 2013 technical report*. Dunedin, New Zealand: Child and Youth Epidemiology Service, University of Otago. Retrieved from <http://www.nzchildren.co.nz/>
- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334. doi:10.1007/BF02310555
- Crosnoe, R., Erickson, K. G., & Dornbusch, S. (2002). Protective functions of family relationships and school factors on the deviant behaviour of adolescent boys and girls: Reducing the impact of risky friendships. *Youth and Society*, *33*(4), 515-544. doi:10.1177/0044118X02033004002
- Cross, A.B., Gottfredson, D.C., Wilson, D.M., Rorie, M., & Connell, N. (2010). Implementation quality and positive experiences in after-school programs. *American Journal of Community Psychology*, *45*(3-4), 370-380. doi:10.1007/s10464-010-9295-z
- Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, *18*(2), 147-167. doi:10.1207/s15327906mbr1802\_2
- Curran, P. J., & Bollen, K. A. (2011). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 107-135). Washington, DC: American Psychological Association.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to non-normality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16-29. doi:10.1037/1082-989X.1.1.16
- Dadds, M. R., Perrin, S., & Yule, W. (1998). Social desirability and self-reported anxiety in children: An analysis of the RCMAS Lie Scale. *Journal of Abnormal Child Psychology*, *26*(4), 311-317. doi:10.1023/A:1022610702439
- Dale, C. M., O'Brien, M., & St. John, S. (Eds.). (2011). *Left further behind: How policies fail the poorest children in New Zealand*. Auckland, New Zealand: Child Poverty Action Group. Retrieved from <http://www.cpag.org.nz/assets/sm/upload/jq/66/v2/dv/WEB%20VERSION%20OF%20LFB.pdf>
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control. *Clinical Psychology Review*, *18*(1), 23-45. doi:10.1016/S0272-7358(97)00043-3
- Danielson, C. K., & Phelps, C. R. (2003). The assessment of children's social skills through self-report: A potential screening instrument for classroom use. *Measurement and Evaluation in Counseling and Development*, *35*(4), 218-229. Retrieved from <http://eric.ed.gov/?id=EJ662361>
- Davey-Smith, G. (Ed.). (2003). *Inequalities in health: Life course perspectives*. Bristol, England: Policy Press.
- Davis, M. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology*, *10*(4), 1-17. Retrieved from [http://www.eckerd.edu/academics/psychology/files/Davis\\_1980.pdf](http://www.eckerd.edu/academics/psychology/files/Davis_1980.pdf)
- DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*(3), 292-307. doi:10.1037/1082-989X.2.3.292

- Deneen, C. C., Brown, G. T. L., Shroff, R. H., & Bond, T. G. (2013). Telling the difference: A first evaluation of an outcome-based learning innovation in teacher education. *Asia-Pacific Journal of Teacher Education*, 41(4), 441-456. doi:10.1080/1359866X.2013.787392
- Denny, S., Peterson, E. R., Stuart, J., Utter, J., Bullen, P., Fleming, T.,...Milfont, T. (2014). Bystander intervention, bullying, and victimization: A multilevel analysis of New Zealand high schools. *Journal of School Violence*, 00, 1-20. doi:10.1080/15388220.2014.910470
- Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (3rd ed., pp. 1-32). Thousand Oaks, CA: Sage.
- DeWit, D. J., Offord, D., Braun, K. (1998). *The relationship between geographic relocation and childhood problem behaviour*. Quebec City, QC: Human Resources Development Canada.
- Ding, C., Liu, Y., & Berkowitz, M. (2011). The study of factor structure and reliability of an abbreviated school climate survey. *Canadian Journal of School Psychology*, 26(3), 241-256. doi:10.1177/0829573511414005
- Dishion, T. J., French, D. C., & Patterson, G. R. (1995). The development and ecology of antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Manual of developmental psychopathology* (pp. 421-471). New York, N.Y: Wiley.
- Ditch, H., & Tarleton, R. (2010, January). *Practical steps: How to tackle variation*. Retrieved from <http://www.educationalleaders.govt.nz/Pedagogy-and-assessment/Evidence-based-leadership/Measuring-learning/How-to-tackle-variation>
- Dixon, P. N., Bobo, M., & Stevick, R. A. (1984). Response differences and preferences for all-category-defined and end-category-defined Likert formats. *Educational and Psychological Measurement*, 44(1), 61-66. doi:10.1177/0013164484441006
- Dobson, J., & Henthorne, K. (1999). *Pupil mobility in schools: Department for education and employment research brief No 168*. London, England: University College London. Retrieved from <http://iris.ucl.ac.uk/iris/publication/10654/1>
- Dobson, J., Henthorne, K., & Lynas, Z. (2000). *Pupil mobility in schools: Final report*. London, England: University College London (Department of Geography, Migration Research Unit).
- Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, 65(2), 649-665. doi:10.2307/1131407
- Donaldson, S. I. (2003). Theory-driven program evaluation in the new millennium. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 109-141). Mahwah, NJ: Lawrence Erlbaum Associates.
- Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. New York, NY: Lawrence Erlbaum Associates.
- Donaldson, J. M., Wiskow, K. M., & Soto, P. L. (2015). Immediate and distal effects of the Good Behaviour Game. *Journal of Applied Behavior Analysis*, 48(3), 685-689. doi: 10.1002/jaba.229
- Doyle, K.A. (1975). *Student evaluation of instruction*. Lexington, MA: Lexington Books.

- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Development, 65*(2), 296-318. doi:10.2307/1131385
- Dunham, T. C., & Davison, M. L. (1991). Effects of scale anchors on student ratings of instructors. *Applied Measurement in Education, 4*(1), 23-35. doi:10.1207/s15324818ame0401\_3
- Durie, M. (1999). Te Pai Mahutonga: A model for Māori health promotion. *Health Promotion Forum of New Zealand Newsletter, 49*, 2-5. Retrieved from <http://www.hauora.co.nz/resources/tepaemahutongatxtvers.pdf>
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3-4), 327-350. doi:10.1007/s10464-008-9165-0
- Dwyer, J., Allison, K., Barrera, M., Hansen, B., Goldenberg, E., & Boutilier, M. (2003). Teachers' perspective on barriers to implementing physical activity curriculum guidelines for school children in Toronto. *Canadian Journal of Public Health, 94*(6), 448-452. Retrieved from [http://www.jstor.org/stable/41993741?seq=1#page\\_scan\\_tab\\_contents](http://www.jstor.org/stable/41993741?seq=1#page_scan_tab_contents)
- Dyson, B., Gordon, B., & Cowan, J. (2011). What is physical education in primary schools in Aotearoa/New Zealand? *Asia-Pacific Journal of Health, Sport and Physical Education, 2*(3-4), 5-16. doi:10.1080/18377122.2011.9730356
- Easton, B., & Ballantyne, S. (2002). *The economic and health status of children*. Wellington, New Zealand: Wellington School of Medicine.
- Eccles, J., & Roeser, R. (1999). School and community influences on human development. In M. Bornstein & M. Lamb (Eds.), *Developmental psychology: An advanced textbook* (4th ed., pp. 503-554). Mahwah, NJ: Erlbaum Press.
- Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., MacIver, D., & Feldaluf, H. (1993). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal, 93*(5), 553-574. doi:10.1086/461740
- Education Counts. (2014). *Number of schools*. Retrieved from [http://www.educationcounts.govt.nz/statistics/schooling/july\\_school\\_roll\\_returns/number-of-schools](http://www.educationcounts.govt.nz/statistics/schooling/july_school_roll_returns/number-of-schools)
- Education Review Office [ERO]. (2007). *Managing transience: Good practice in primary schools*. Wellington, New Zealand: Education Review Office. Retrieved from <http://www.ero.govt.nz/National-Reports/Managing-Transience-Good-Practice-in-Primary-Schools-April-2007/Introduction>
- Edwards, J. (1997). *Student transience: Issues for schools arising from school mobility*. (Unpublished paper). Adelaide, Australia: University of South Australia.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.
- Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Hillsdale, NJ: Lawrence Erlbaum.
- Eisenberg, N., Fabes, R. A., Murphy, B. C., Karbon, M., Smith, M., & Maszk, P. (1996). The relations of children's dispositional empathy-related responding to their emotionality, regulation, and social functioning. *Developmental Psychology, 32*(2), 195-209. doi:10.1037/0012-1649.32.2.195
- Elder, G. H., Van Nguyen, T., & Caspi, A. (1995). Linking family hardship to children's lives. *Child Development, 56*(2), 361-375. doi:10.2307/1129726

- Engels, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 196(4286), 129-136. doi:10.1126/science.847460
- Eriksson, M. & Lindstrom, B. (2007). Antonovsky's sense of coherence scale and its relation with quality of life: A systematic review. *Journal of Epidemiology and Community Health*, 61(11), 938-944. doi:10.1136/jech.2006.056028
- Evans, A. N. (1992). *Using basic statistics in the social sciences* (3rd ed.). Scarborough, ON: Prentice Hall Allyn and Bacon Canada.
- Evans, A.N. (2014). *Using basic statistics in the behavioral and social sciences* (5th ed.). Edmonton, AB: Sage.
- Ewing, J. L. (1970). *Development of the New Zealand primary school curriculum 1877 – 1970*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Expert Advisory Group on Solutions to Child Poverty [EAG]. (2012). *Solutions to child poverty in New Zealand: Evidence for action*. Wellington, New Zealand: Office of the Children's Commissioner. Retrieved from [http://www.occ.org.nz/publications/child\\_poverty](http://www.occ.org.nz/publications/child_poverty)
- Expert Advisory Panel on Solutions to Child Poverty [EAG]. (2014). *Measuring child poverty: Introduction*. Wellington, New Zealand: Ministry of Social Development. Retrieved from [http://www.nzchildren.co.nz/document\\_downloads/B.%202014%20CPM%20Child%20Poverty%20and%20Living%20Standards.pdf](http://www.nzchildren.co.nz/document_downloads/B.%202014%20CPM%20Child%20Poverty%20and%20Living%20Standards.pdf)
- Fan, X., Miller, B. C., Park, K.-E., Winward, B. W., Christensen, M., Grotevant, H. D., & Tai, R. H. (2006). An exploratory study about inaccuracy and invalidity in adolescent self-report surveys. *Field Methods*, 18(3), 223-244. doi:10.1177/152822X06289161
- Fan, X. & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 42(3), 509-529. doi:10.1080/00273170701382864
- Felner, R., & Adam, A. (1988). The school transitional environment project: An ecological intervention and evaluation. In R. Price, E. Cowen, R. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners* (pp. 111-122). Washington, DC: American Psychological Association.
- Field, A. (2013a). Chapter 11: Comparing several means. In A. Field (Ed.), *Discovering statistics using IBM SPSS statistics* (4th ed., pp. 429-477). London, England: Sage.
- Field, A. (2013b). *Discovering statistics using IBM SPSS Statistics* (4th ed). London, England: Sage.
- Fields, B. A. (1997). Children on the move: The social and educational effects of family mobility. *Children Australia*, 22(3), 4-9. Retrieved from <http://search.informit.com.au/fullText;dn=980303116;res=APAFT>.
- Finkel, S. (1995). *Causal analysis with panel data*. Thousand Oaks, CA: Sage.
- Finn, R. H. (1972). Effects of some variations in rating scale characteristics on the means and reliabilities of ratings. *Educational and Psychological Measurement*, 32(2), 255-265. doi: 10.1177/001316447203200203
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. D. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269-314). Greenwich, CT: Information Age Publishing.

- Fisher, T.A., Matthews, L., Stafford, M.E., Nakagawa, K., Durante, K. (2002). School personnel's perceptions of effective programs for working with mobile students and families. *The Elementary School Journal*, 102(4), 317-333. doi:10.1086/499706
- Foundation for Youth Development [FYD]. (2012). *The Kiwi Can co-ordinator manual*. Auckland, New Zealand: The Foundation for Youth Development.
- Foundation for Youth Development [FYD]. (2014). *FYD – Who we are*. Retrieved from <http://www.fyd.org.nz/about/who-we-are>
- Fraser, B. J. (1980). Research on classroom learning environment in the 1970's and 1980's. *Studies in Educational Evaluation*, 6(3), 221-223. doi:10.1016/0191-491X(80)90025-5
- Fraser, B. J. (1981). Learning environment in curriculum evaluation: A review. *Evaluation in Education*, 5(1), 1-93. doi:10.1016/0191-765X(81)90014-8
- Fraser, B. J. (1982). Development of short forms of several classroom environment scales. *Journal of Educational Measurement*, 19(3), 221-227. doi:10.1111/j.1745-3984.1982.tb00130.x
- Fraser, B. J., Anderson, G. J., & Walberg, H. J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory and My Class Inventory*. Perth, Australia: Western Australian Institute of Technology.
- Fraser, B. J., & Fisher, D. L. (1986). Using short forms of classroom climate instruments to assess and improve classroom psychosocial environments. *Journal of Research in Science Teaching*, 23(5), 387-413. doi:10.1002/tea.3660230503
- Frisbie, D. A., & Brandenburg, D. C. (1979). Equivalence of questionnaire items with varying response formats. *Journal of Educational Measurement*, 16(1), 43-48. doi:10.1111/j.1745-3984.1979.tb00085.x
- Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings*. Boston, MA: Kluwer Academic.
- Garbarino, J. (1985). *Adolescent development: An ecological perspective*. Columbus, OH: Charles E. Merrill Publishing.
- Garrett, P., Ng'andu, N., & Ferron, J. (1994). Poverty experiences of young children and the quality of their home environments. *Child Development*, 65(2), 331-345. doi:10.2307/1131387
- Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 40-65). Newbury Park, CA: Sage.
- Gestsdottir, S., & Lerner, R. M. (2007). Intentional self-regulation and positive youth development in early adolescence: Findings from the 4-H study of positive youth development. *Developmental Psychology*, 43(2), 508-521. doi:10.1037/0012-1649.43.2.508
- Gibbons, S. & Telhaj, S. (2011). Pupil mobility and school disruption. *Journal of Public Economics*, 95(9-10), 1156-1167. doi:10.1016/j.jpubeco.2011.03.004
- Gilbert, J. (2005). *Educational issues for communities affected by transience and residential mobility: Report on Phase I (2003-2004)*. Wellington, New Zealand: New Zealand Council for Educational Research.

- Gilbert, J. & Bull, A. (2007). *Educational implications for communities affected by transience and residential movement*. Wellington, New Zealand: New Zealand Council for Educational Research. Retrieved from [http://www.google.co.nz/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCMQFjAA&url=http%3A%2F%2Fwww.cresa.co.nz%2Fwp-content%2Fuploads%2F2011%2F06%2Fmobility-and-educationfinal-\\_gilbert-and-bull\\_.pdf&ei=qj0WVOajMZKwuATSuICwBg&usq=AFQjCNFjZolOdtwDtexrXVvyEbhc-NHRxQ&sig2=xBUyO0CpxP-I\\_85hZW6hDg&bvm=bv.75097201,d.c2E](http://www.google.co.nz/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCMQFjAA&url=http%3A%2F%2Fwww.cresa.co.nz%2Fwp-content%2Fuploads%2F2011%2F06%2Fmobility-and-educationfinal-_gilbert-and-bull_.pdf&ei=qj0WVOajMZKwuATSuICwBg&usq=AFQjCNFjZolOdtwDtexrXVvyEbhc-NHRxQ&sig2=xBUyO0CpxP-I_85hZW6hDg&bvm=bv.75097201,d.c2E)
- Gottfredson, D.C., & Gottfredson, G.D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency*, 39(1), 3-36. doi:10.1177/002242780203900101
- Government of Canada. (2007). *Poverty profile: Special edition*. Ottawa, ON: National Council of Welfare. Retrieved from [http://www.esdc.gc.ca/eng/communities/reports/poverty\\_profile/snapshot.shtml](http://www.esdc.gc.ca/eng/communities/reports/poverty_profile/snapshot.shtml)
- Gower, J. C. (1967). A comparison of some methods of cluster analysis. *Biometrics*, 23(4), 623-637. doi:10.2307/2528417
- Gradin, C. (2008). *Poverty among minorities in the United States: Explaining racial poverty gap for Blacks and Latinos*. Ithaca, NY: Society for the Study of Economic Inequality. Retrieved from <http://www.ecineq.org/milano/WP/ECINEQ2008-96.pdf>
- Grant, J. P. (Ed.). (1994). *The progress of nations - Unicef*. Oxford, England: Oxford University Press. <http://www.unicef.org/sowc/archive/ENGLISH/The%20State%20of%20the%20World%27s%20Children%201994.pdf>
- Griffith, J. (2000). School climate as group evaluation and group consensus: Student and parent perceptions of the school environment. *Elementary School Journal*, 101(1), 35-61. doi:10.1086/499658
- Griggs, G. (2008). Outsiders inside: The use of sports coaches in primary schools. *Physical Education Matters Spring*, 3(2), 33-36. Retrieved from [http://www.academia.edu/4524067/Outsiders\\_inside\\_The\\_use\\_of\\_sports\\_coaches\\_in\\_primary\\_schools](http://www.academia.edu/4524067/Outsiders_inside_The_use_of_sports_coaches_in_primary_schools)
- Griggs, G. (2010). For sale – Primary school physical education. £20 per hour or nearest offer. *Education*, 38(1), 39-46. doi:10.1080/03004270903099793
- Hall, I. M., & Hall, D. J. (2004). *Evaluation and social research: Introducing small-scale practice*. New York, NY: Palgrave Macmillan.
- Hancock, G. R., & Klockars, A. J. (1991). The effect of scale manipulations on validity: Targeting frequency rating scales for anticipated performance levels. *Applied Ergonomics*, 22(3), 147-154. doi:10.1016/0003-6870(91)90153-9
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, 88(9-10), 1721-1746. doi:10.1016/S0047-2727(03)00063-X
- Harter, S. (1982). The perceived competence scale for children. *Child Development*, 53(1), 87-98. doi:10.2307/1129640
- Hartman, C. (2002). High classroom turnover: How children get left behind. In D. M. Piche, W.L. Taylor & R. A. Reed (Eds.), *Citizens' commission on human rights, part two: Discriminatory practices in education* (pp. 227-244). Washington, DC: Citizens Commission on Civil Rights. Retrieved from <http://files.eric.ed.gov/fulltext/ED463386.pdf>

- Hawkins, J. D., Kosterman, R., Catalano, R. F., Hill, K. G., & Abbott, R. D. (2005). Promoting positive adult functioning through social development intervention in childhood: Long-term effects from the Seattle Social Development Project. *Archives of Pediatrics & Adolescent Medicine*, *159*(1), 25-31. doi: 10.1001/archpedi.159.1.25
- Hayes, A. F., & Kai, L. (2007). Using heteroskedasticity-consistent standard error estimates in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*(4), 709-722. doi:10.3758/BF03192961
- Hays, K. (2000). *Investigation report into transient children*. [Unpublished paper]. Christchurch, New Zealand: Christchurch College of Education.
- Heck, K. E., & Subramaniam, A. (2009). *Youth development frameworks*. [Monograph]. Davis, CA: University of California.
- Helu, S. L., Robinson, E., Grant, S., Herd, R., & Denny, S. (2009). *Youth '07: The health and wellbeing of secondary school students in New Zealand: Results for Pacific young people*. Auckland, New Zealand: The University of Auckland.
- Hendricks Brown, C., Kellman, S., Ialongo, N., Poduska, J., & Ford, C. (2007). Prevention of aggressive behaviour through middle-school using a first-grade classroom-based intervention. In M. Tsuang and M. Lyons (Eds.). *Recognition and prevention of major mental and substance abuse disorders* (pp. 347-369). Arlington, VA: American Psychiatric Publishing.
- Henshaw, A., Kelly, J., & Gratton, J. (1992). Skipping's for girls: children's perceptions of gender roles and gender preferences. *Educational Research*, *34*(3), 229-235. doi:10.1080/0013188920340307
- Herrenkohl, T. I., Jengeun, L., & Hawkins, D. (2012). Risk versus direct protective factors and youth violence. *American Journal of Preventive Medicine*, *43*(2S1), S41-S56. doi: 10.1016/j.amepre.2012.04.030
- Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging*, *18*(4), 639-657. doi: 10.1037/0882-7974.18.4.639
- Hill, A., Navin, F., & Lynch, A. (2009). *Coming to grips with student mobility and policy implications: A case study from regional Queensland*. Paper presented at AARE 2009 International Education Research Conference, Macquarie University, Canberra, Australia. Retrieved from <http://researchonline.jcu.edu.au/18483/1/hil091540.pdf>
- Hodges, E. V. E., Boivin, M., Vitaro, F., & Bukowski, W. M. (1999). The power of friendship: Protection against an escalating cycle of peer victimization. *Developmental Psychology*, *35*(1), 94-101. doi:10.1037/0012-1649.35.1.94
- Hopkins, K. D., Hopkins, B. R., & Glass, G. V. (1996). *Basic statistics for the behavioural sciences* (3rd ed.). Boston, MA: Allyn & Bacon.
- House, E. R. (1993). *Professional evaluation, social impact and political consequences*. Newbury Park, CA: Sage.
- Howard-Barr, E. M., Rienzo, B. A., Pigg, M. R., & James, D. (2005). Teacher beliefs, professional preparation, and practices regarding exceptional students and sexuality education. *Journal of School Health*, *75*(3), 99-104. doi:10.1111/j.1746-1561.2005.tb06649.x
- Hu, L. T. & Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to under-parameterized model misspecification. *Psychological Methods*, *3*(4), 424-453. doi:10.1037/1082-989X.3.4.424

- Hu, L. T. & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. doi:10.1080/10705519909540118
- IBM. (2011). *Structural equation modelling with IBM SPSS Amos*. Retrieved from <http://www-03.ibm.com/software/products/en/spss-amos/>
- Innocenti Research Centre. (2012). *Measuring child poverty: Report card 10*. Florence, Italy: UNICEF Innocenti Research Centre. Retrieved from [http://www.unicef-irc.org/publications/pdf/rc10\\_eng.pdf](http://www.unicef-irc.org/publications/pdf/rc10_eng.pdf)
- Iyer, P. & Aggleton, P. (2013) 'Sex education should be taught, fine...but we make sure they control themselves': Teachers' beliefs and attitudes towards young people's sexual and reproductive health in a Ugandan secondary school. *Sex Education: Sexuality, Society and Learning*, 13(1), 40-53. doi:10.1080/14681811.2012.677184
- Jarrett, R. L. (1997). Bringing families back in: Neighborhoods' effects on child development. In J. Brooks-Gunn, G. J. Duncan, & J. L. Aber (Eds.), *Neighborhood poverty: Policy implications in studying neighborhoods* (Vol. 2, pp. 48-64). New York, NY: Russell Sage Foundation.
- Jelicic, H., Bobek, D. L., Phelps, E., Lerner, R. M., & Lerner J. V. (2007). Using positive youth development to predict contribution and risk behaviours in early adolescence: Findings from the first two waves of the 4-H Study of Positive Youth Development. *International Journal of Behavioural Development*, 31(3), 263-273. doi:10.1177/0165025407076439
- Jelleyman, T., & Spencer, N. (2008). Residential mobility in childhood and health outcomes: A systematic review. *Journal of Epidemiology and Community Health*, 62(7), 584-592. doi: 10.1136/jech.2007.060103
- Jenkins, Jr., G. D., & Taber, T. D. (1977). A Monte-Carlo study of factors affecting three indices of composite scale reliability. *Journal of Applied Psychology*, 62(4), 392-398. doi:10.1037/0021-9010.62.4.392
- Johnson, A. (2002). *A study of student transience in South Auckland primary schools*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 409-426. doi:10.1007/BF02291366
- Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal development investigations. In J. R. Nesselroade and P. B. Baltes (Eds.), *Longitudinal research in the study of behaviour and development* (pp. 303-352). New York, NY: Academic Press.
- Jöreskog, K. G., & Sorbom, D. (1981). *LISREL VI user's guide* (3rd ed.). Mooresville, IN: Scientific Software.
- Jöreskog, K. G., & Sorbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Lisrel 8: Lawrence Erlbaum Associates.
- Jourdain, D., Pommier, J., & Quidu, F. (2010). Practices and representations of health education among primary school teachers. *Scandinavian Journal of Public Health*, 38(1), 86-94. doi:10.1177/1403494809350518
- Kariuji, P., & Nash, J. (1999, November). *The relationship between multiple school transfers during elementary years and student achievement*. Paper presented at the annual conference of the Mid-South Educational Research Association, Point Clear, AL.
- Karlsson, R., & Rastam, A. S. (1985). *Halsofostran i Skolan: Mal coh Verklighet*. Hogskolan I, Jonkoping: Institutionen for Beteendevetenskap.

- Kawachi, I., Kennedy, B. P., & Glass, R. (1999). Social capital and self-rated health: A contextual analysis. *American Journal of Public Health, 89*(8), 1187-1193. doi:10.2105/AJPH.89.8.1187
- Keenan, M., Moore, J.L., & Dillenburger, K. (2000). The Good Behaviour Game. *Child Care in Practice, 6*(1), 27-38. doi: 10.1080/135270008413191
- Kelsey, J. (1998). *The New Zealand experiment*. Auckland, New Zealand: Auckland University Press.
- Kenny, D. A. (1975a). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin, 82*(6), 887-903. doi:10.1037/0033-2909.82.6.887
- Kenny, D. A. (1975b). A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design. *Psychological Bulletin, 82*(3), 345-362. doi:10.1037/0033-2909.82.6.887
- Kenny, D. A., & Harachiewicz, J. M. (1979). Cross-lagged panel correlation: Practice and promise. *Journal of Applied Psychology, 64*(4), 372-379. doi:10.1037/0033-2909.82.3.345
- Kerbow, D. (1996). Patterns of urban student mobility and local school reform. *Journal of Education from Students Placed at Risk, 1*(2), 147-169. doi:10.1080/00220671.1996.9944438
- Khoury-Kassabri, M., Benbenishty, R., Astor, R. A., & Zeira, A. (2004). The contributions of community, family, and school variables to student victimization. *American Journal of Community Psychology, 34*(3/4), 187-204. doi:10.1007/s10464-004-7414-4
- Kincaid, D., Childs, K., Blasé, K.A., & Wallace, F. (2007). Identifying barriers and facilitators in implementing schoolwide positive behavior support. *Journal of Positive Behavior Interventions, 9*(3), 174-184. doi:10.1177/10983007070090030501
- Klesges, L. M., Baranowski, T., Beech, B., Cullen, K., Murray, D. M., Rochon, J., & Pratt, C. (2004). Social desirability bias in self-reported dietary, physical activity and weight concerns measures in 8-to10-year old African-American girls: Results from the Girls Health Enrichment Multisite Studies (GEMS). *Preventive Medicine, 38*, S78-S87. doi:10.1016/j.ypmed.2003.07.003
- Kline, P. (1993). *The handbook of psychological testing*. London, England: Routledge.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.
- Korbin, J. E. (1992). Child poverty in the United States. *American Behavioral Scientist, 35*(3), 213-219. doi:10.1177/0002764292035003002
- Korenman, S., Miller, J.E., & Sjaastad, J.E. (1995). Long-term poverty and child development in the United States: Results from the NLSY. *Child Youth Service Review, 17*(1-2), 127-155. doi: 10.1016/0190-7409(95)00006-X
- Koschmider, S. M. S. (1997). *Transience and education: Academic and psychosocial effects of geographic mobility on children*. (Unpublished masters' thesis). University of Toronto, Toronto, Canada). Retrieved from <https://tspace.library.utoronto.ca/bitstream/1807/11599/1/MQ29160.pdf>
- Kuperminc, G., Leadbeater, B. J., Emmons, C., & Blatt, S. J. (1997). Perceived school climate and problem behaviors in middle school students: The protective function of a positive educational environment. *Journal of Applied Developmental Science, 1*(2), 76-88. doi:10.1207/s1532480xads0102\_2

- Kupersmidt, J. B., Burchinal, M., & Patterson, C. J. (1995). Developmental patterns of childhood peer relations as predictors of externalizing behavior problems. *Development and Psychopathology*, 7(4), 825–843. doi:10.1017/S0954579400006866
- Ladd, G., & Price, J. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development*, 58(5), 1168-1189. doi:10.2307/1130613
- Lai, J. C. L. (1994). Differential predictive power of the positively versus the negatively worded items of the Life Orientation Test. *Psychological Reports*, 75(3f), 1507-1515. doi:10.2466/pr0.1994.75.3f.1507
- Lalonde M. (1974). *A new perspective on the health of Canadians*. Ottawa, ON: Information Canada.
- Lam, T. C. M., & Klockars, A. J. (1982). Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement*, 19(4), 317-322. doi:10.1111/j.1745-3984.1982.tb00137.x
- Lang, K. (2002). *Measuring ethnicity in the New Zealand census*. Wellington, New Zealand: Statistics New Zealand.
- Larson, R. W., & Brown, J. R. (2007). Emotional development in adolescence: What can be learned from a high school theater program? *Child Development*, 78(4), 1083-1099. doi: 10.1111/j.1467-8624.2007.01054.x
- Lavin, J., Swindlehurst, G., & Foster, V. (2008). The use of coaches, adults supporting learning and teaching assistants in the teaching of physical education in the primary school. *Physical Education Matters Spring*, 3(1), ix–xi. Retrieved from <http://www.jimlavin.co.uk/publications/>
- Lee, A. (2000). *Transient children: Perceptions of how often transient children come and go*. (Unpublished master's thesis). Auckland College of Education, Auckland, New Zealand.
- Leffert, N., Benson, P.L., Scales, P. C., Sharma, A. R., Drake, D. R., & Blyth, D. A. (1998). Developmental assets: Measurement and prediction of risk behaviour among adolescents. *Applied Developmental Science*, 2(4), 209-230. doi:10.1207/s1532480xads0204\_4
- Lerner, R. M. (2004). *Liberty: Thriving and civic engagement among American youth*. Thousand Oaks, CA: Sage.
- Lerner, R. M. (2007). *The good teen: Rescuing adolescents from the myths of the storm and stress years*. New York, NY: The Crown Publishing Group.
- Lerner, R. M., Almergi, J. B., Theokas, C. & Lerner, J. V. (2005). Positive youth development: A view of the issues. *Journal of Early Adolescence*, 25(1), 10-16. doi:10.1177/0272431604273211
- Lerner, R. M., Brentano, C., Dowling, E. M., & Anderson, P. M. (2002). Positive youth development: Thriving as the basis of personhood and civil society. *New Directions for Youth Development*, 95(11), 11-33. doi:10.1002/yd.14
- Lerner, R. M., & Castellino, D. R. (2002). Contemporary developmental theory and adolescence: Developmental systems and applied developmental science. *Journal of Adolescent Health*, 31(6), 122–135. doi:10.1016/S1054-139X(02)00495-0

- Lerner, R. M., Lerner, J. V., Almerigi, J., Theokas, C., Phelps, E., Gestsdottir, S., & von Eye, A. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth grade adolescents: Findings from the first wave of the 4-H study of positive development. *Journal of Early Adolescence*, 25(1), 17-71. doi:10.1177/0272431604272461
- Lerner, R. M., von Eye, A., Lerner, J. V., Lewin-Bizan, S., & Bowers, E. P. (2010). The meaning and measurement of thriving: A view of the issues. *Journal of Youth and Adolescence*, 39(7), 707-719. doi:10.1007/s10964-010-9531-8
- Lichter, D. T., & Jensen, L. I. (2002). Rural America in transition: Poverty and welfare at the turn of the 20th century. In L. A. Whitener, B. A. Weber, & G. Duncan (Eds.), *Rural dimensions of welfare reform* (pp. 77–112). Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- Liem, R., & Liem, J. (1978). Social class and mental illness reconsidered: The role of economic stress and social support. *Journal of Health and Social Behavior*, 19(2), 139-156. doi:10.2307/2136530
- Lindstrom, B., & Eriksson, M. (2006). Contextualizing salutogenesis and Antonovsky in public health development. *Health Promotion International*, 21(3), 238-244. doi:10.1093/heapro/dal016
- Lipsey, M. W., & Cordray, D. S. (2000). Evaluation methods for social intervention. *Annual Review of Psychology*, 51(1), 345-375. doi:10.1146/annurev.psych.51.1.345
- Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte-Carlo approach. *Journal of Applied Psychology*, 60(1), 10-13. doi:10.1037/h0076268
- Lloyd, G., Stead, J., Jordan, E., & Norris, C. (1998). Teachers and Gypsy Travellers. *Scottish Education Review*, 3(1), 48-65. Retrieved from <http://www.scotedreview.org.uk/pdf/65.pdf>
- Lockwood, V. (2002). Poverty in paradise: Development and relative income poverty in rural Tahitian society. *Human Organization*, 61(3), 210-225. Retrieved from <https://workfamily.sas.upenn.edu/wfrn-repo/object/9mj9q064bm89xq2u>
- Loeber, R., & Stouthamer-Loeber, M. (1986). Family factors as correlates and predictors of juvenile conduct problems and delinquency. In M. Tonry & N. Morris (Eds.). *Crime and Justice* (pp. 29-149). Chicago, IL: University of Chicago Press.
- Lonczak, H. S., Abbott, R.D., Hawkins, J. D., Kosterman, R., Catalano, R.F. (2002). Effects of the Seattle social development project on sexual behavior, pregnancy, birth, and sexually transmitted disease outcomes by age 21 years. *Archives of Pediatrics & Adolescent Medicine*, 156(5), 438-447. doi: 10.1001/archpedi.156.5.438
- Low, J., & Jose, P. (2011). *Lifespan development: New Zealand perspectives* (2nd ed.). Auckland, New Zealand: Pearson Education New Zealand.
- Lynch, T. (2007). What has changed since the 1992 Senate inquiry into physical and sport education? An evaluation of school responses within three Brisbane Catholic Education (BCE) primary schools. *ACHPER Australia Healthy Lifestyles Journal*, 54(1), 16–23. doi:10.1016/j.ijer.2015.022.003
- Lytle, L., Ward, J., Nader, P., Pedersen, S., & Williston, B. (2003). Maintenance of a health promotion program in elementary schools: results from the CATCH-ON study key informant interviews. *Health Education and Behavior*, 30(4), 503–518. doi:10.1177/1090198103253655

- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on change. *Psychological Bulletin*, *111*(3), 490-504. doi:10.1037/0033-2909.111.3.490
- Mare, R.D. (1982). Socio-economic status and child mortality. *American Journal of Public Health*, *72*(6), 539-547. doi: 10.2105/AJPH.72.6.539
- Mark, M. M. (2003). Toward an integrative view of the theory and practice of program and policy evaluation. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millenium* (pp. 183-204). Mahwah, NJ: Lawrence Erlbaum Associates.
- Marmot, M. G., & Wilkinson, R. G. (2006). *Social determinants of health* (2nd ed.). Oxford, England: Oxford University Press. Retrieved from <http://www.oxfordscholarship.com.ezproxy.auckland.ac.nz/view/10.1093/acprof:oso/9780198565895.001.0001/acprof-9780198565895>
- Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, *6*(4), 311-360. doi:10.1207/s15327574ijt0604\_1
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cut-off values for fit indices and dangers in overgeneralizing Hu & Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320-341. doi:10.1207/s15328007sem1103\_2
- Mazor, K. M., Clauser, B. E., Field, T., Yood, R. A., & Gurwitz, J. H. (2002). A demonstration of the impact of response bias on the results of patient satisfaction surveys. *Health Services Research*, *37*(5), 1403-1417. doi:10.1111/1475-6773.11194
- McDowell, I. & Newell, C. (1996). *Measuring health: A guide to rating scales and questionnaires* (2<sup>nd</sup> ed.). New York, NY: Oxford University Press.
- McGee, C., Harlow, A., Miller, T., Cowie, B., Hill, M., Jones, A., & Donaghy, A. (2004). *Teachers' experiences in curriculum implementation: General curriculum, the arts and health and physical education* (pp. 295 - 309). Wellington, New Zealand: Ministry of Education.
- McKelvie, S. G. (1978). Graphic rating scales: How many categories? *British Journal of Psychology*, *69*(2), 185-202. doi:10.1111/j.2044-8295.1978.tb01647.x
- McLeod, E. (2014, April). The terrible toll of transience on our schools and students. *Education Aotearoa*, (p.1). Retrieved from <http://www.educationaotearoa.org.nz/all-stories/2014/4/28/the-terrible-toll-of-transience-on-our-schools-and-students.html>
- McLoyd, V. C., Jayarante, T. E., Ceballo, R., & Borquez, J. (1994). Unemployment and work interruption among African American single mothers: Effects on parenting and adolescent socioemotional functioning. *Child Development*, *65*(2), 562-589. doi:10.2307/1131402
- Menaghan, E. G. (1999). Social stressors in childhood and adolescence. In A. V. Horwitz & T. L. Scheid (Eds.), *A handbook for the study of mental health: Social contexts, theories, and systems* (pp. 315-327). New York, NY: Cambridge University Press.
- Merry, S. (1981). *Urban danger: Life in a neighbourhood of strangers*. Philadelphia, PA: Temple University Press.
- Mertens, D. M. & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. New York, NY: Guilford Press.

- Midgley, C., & Edelin, K. C. (1998). Middle school reform and early adolescent wellbeing: The good news and the bad. *Educational Psychologist, 33*(4), 195–206. doi:10.1207/s15326985ep3304\_4
- Milligan, G. W. (1980). An examination of the effect of six types of error perturbation of fifteen clustering algorithms. *Psychometrika, 45*(3), 325-342. doi:10.1007/BF02293907
- Ministry of Education [MOE]. (1999). *Health and physical education in the New Zealand curriculum*. Wellington, New Zealand: Learning Media.
- Ministry of Education [MOE]. (2007). *The New Zealand curriculum for English-medium teaching and learning in years 1-13*. Wellington, New Zealand: The New Zealand Government.
- Ministry of Education [MOE]. (2013a). *Education counts – Transient students*. Retrieved from <http://www.educationcounts.govt.nz/indicators/main/student-engagement-participation/transient-students>
- Ministry of Education [MOE]. (2014a). *New Zealand education system overview*. Retrieved from [http://www.minedu.govt.nz/NZEducation/EducationPolicies/InternationalEducation/ForInternationalStudentsAndParents/NZEdOverview/Education\\_System.aspx](http://www.minedu.govt.nz/NZEducation/EducationPolicies/InternationalEducation/ForInternationalStudentsAndParents/NZEdOverview/Education_System.aspx)
- Ministry of Education [MOE]. (2014b). *School decile ratings*. Retrieved from <http://www.minedu.govt.nz/Parents/AllAges/EducationInNZ/SchoolsInNewZealand/SchoolDecileRatings.aspx>
- Ministry of Education [MOE]. (2014c). *Decile ratings*. Retrieved from <http://www.minedu.govt.nz/NZEducation/EducationPolicies/Schools/SchoolOperations/Resourcing/ResourcingHandbook/Chapter1/DecileRatings.aspx>
- Ministry of Health [MOH]. (2012). *The health of New Zealand children 2011/12: Key findings of the New Zealand health survey*. Wellington, New Zealand: Ministry of Health.
- Ministry of Health [MOH]. (2013). *New Zealand health survey: Annual update of key findings 2012/13*. Wellington, New Zealand: Ministry of Health.
- Moos, R. H. (1974). *The social climate scales: An overview*. Palo Alto, CA: Consulting Psychologists Press.
- Morgan, P., & Hansen, V. (2007). Recommendations to improve primary school physical education: Classroom teachers' perspective. *The Journal of Educational Research, 101*(2), 99-111. doi:10.3200/JOER.101.2.99-112
- Morgan, P., & Hansen, V. (2008). Classroom teachers' perceptions of the impact of barriers to teaching physical education on the quality of physical education programs. *Research Quarterly for Exercise and Sport, 79*(4), 506-516. doi:10.1080/02701367.2008.10599517
- Mullen, P. D., Hersey, J. C., & Iverson, D. C. (1987). Health behaviour models compared. *Social Science Medicine, 24*(11), 973-981. doi:10.1016/0277-9536(87)90291-7
- Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology, 25*(1), 3–53. doi:10.1006/ceps.1999.1019
- Murrow, K., Kalafatelis, M., Fryer, M., Ryan, N., & Dowden, A. (2002). *An evaluation of three programmes in the innovations funding pool: Kiwi Can*. Wellington, New Zealand: Ministry of Education.

- Nastasi, B. K. & Hitchcock, J. (2009). Challenges of evaluating multilevel interventions. *American Journal of Community Psychology*, 43(3-4), 360-376. doi:10.1007/s10464-009-9239-7
- Neighbour, M. (2000). *Developing an understanding of transient pupils*. Whangarei, New Zealand: Tai Tokerau Education Centre.
- Neighbour, M. (2002). *New Zealand principals' federation transient pupil survey 2001-2002*. Wellington, New Zealand: New Zealand Principals' Federation. Retrieved from <http://www.nzpf.ac.nz/>
- Neighbour, M. (2003, June). Transient children: What works best? *New Zealand Principal*. Retrieved from [http://www.nzpf.ac.nz/sites/default/files/June03\\_Transcience.pdf](http://www.nzpf.ac.nz/sites/default/files/June03_Transcience.pdf)
- Newman, I., & Hitchcock, J. (2011). Underlying agreements between quantitative and qualitative research: The short and tall of it all. *Human Resource Development Review*, 10(4), 381-398. doi:10.1177/1534484311413867
- New Zealand Council for Educational Research. (2005). *Educational issues for communities affected by transience and residential mobility*. Wellington, New Zealand: New Zealand Council for Educational Research.
- New Zealand Education Department [NZED]. (1928). *Syllabus of instructions for public schools*. Wellington, New Zealand: New Zealand Education Department.
- New Zealand Education Department [NZED]. (1935). *Syllabus of instructions for public schools*. Wellington, New Zealand: New Zealand Education Department.
- New Zealand Government. (2012a). *The green paper for vulnerable children: Every child thrives, belongs and achieves*. Wellington, New Zealand: Ministry of Social Development and Employment.
- New Zealand Government. (2012b). *Children's action plan: Identifying, supporting and protecting vulnerable children - The White Paper for vulnerable children*. Wellington, New Zealand: Ministry of Social Development and Employment. Retrieved from <http://www.msd.govt.nz/documents/about-msd-and-our-work/work-programmes/policy-development/white-paper-vulnerable-children/white-paper-for-vulnerable-children-childrens-action-plan-summaries.pdf>
- New Zealand Peace Foundation. (2014). *Cool Schools peer mediation programme homepage*. Retrieved on from <http://www.peace.net.nz/index.php?pageID=38>
- Nutbeam, D. (1998). *Health promotion glossary*. Geneva, Switzerland: World Health Organization.
- O'Brien, V., Martinez-Pons, M., & Kopala, M. (1999). Mathematics self-efficacy, ethnic identity, gender, and career interests related to mathematics and science. *Journal of Educational Research*, 92(4), 231-235. doi:10.1080/00220679909597600
- Organisation for Economic Cooperation and Development [OECD]. (2009). Chapter 2: Comparative child wellbeing across the OECD. In OECD, (Ed.), *Doing better for children* (pp.21-63). Retrieved from <http://www.oecd.org/els/family/43570328.pdf>
- Ory, J. C., & Wise, S. L. (1981, April). *Attitude change measured by scales with 4 and 5 response options*. Paper presented at the annual meeting of the National Council on Measurement in Education, Northern Illinois University, Chicago, IL.

- Ottewill, R., & Wall, A. (2004). *Models of health and health care*. London, England: Crown Copyright. Retrieved from [https://www.ihm.org.uk/documents/Education/mhsc\\_resource\\_file/diploma\\_managing\\_your\\_enterprise/Diploma\\_managing\\_your\\_enterprise\\_models\\_of\\_health\\_and\\_health\\_care](https://www.ihm.org.uk/documents/Education/mhsc_resource_file/diploma_managing_your_enterprise/Diploma_managing_your_enterprise_models_of_health_and_health_care)
- Otunuku, M., & Brown, G. T. L. (2007). Tongan students' attitudes towards their subjects in New Zealand relative to their academic achievement. *Asia Pacific Education Review*, 8(1), 117-128. doi:10.1007/BF03025838
- Ovens, A., & Cassidy, T. (2009). Curriculum acoustics: Analysing the changing voice of the New Zealand health and physical education curriculum. In M. Dinan Thompson (Ed.), *Health and physical education: Issues for curriculum in Australia and New Zealand* (pp. 22-37). Wellington, New Zealand: Oxford University Press.
- Pateman, B., Grunbaum, J., & Kann, L. (1995). Voices from the field – A qualitative analysis of classroom, school, district, and state health education policies and programs. *Journal of School Health*, 69(7), 258-262. doi:10.1111/j.1746-1561.1999.tb06405.x
- Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice*, 15(3), 311-319. doi:10.1016/0886-1633(94)90026-4
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Payne, A. A., Gottfredson, D. C., & Gottfredson, G. D. (2003). Schools as communities: The relationships among communal school organization, student bonding, and school disorder. *Criminology*, 41(3), 749-777. doi:10.1111/j.1745-9125.2003.tb01003.x
- Penney, D. (2008). Playing a political game and playing for position: Policy and curriculum development in health and PE. *European Physical Education Review*, 4(1), 33-49. doi: 10.1177/1356336X07085708
- Perry, B. (2010). *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982-2009*. Wellington, New Zealand: Ministry of Social Development.
- Perry, B. (2011). *Household incomes in New Zealand: Trends in indicators of inequality and hardship*. Wellington, New Zealand: Ministry of Social Development.
- Perry B. (2013). *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982 to 2012*. Wellington, New Zealand: Ministry of Social Development. Retrieved from <http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/monitoring/household-incomes/>
- Perry, B. (2014). *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982 to 2013*. Wellington, New Zealand: Ministry of Social Development. Retrieved from <http://www.msd.govt.nz/about-msd-and-our-work/publicationsresources/monitoring/household-incomes/index.html>
- Peterson, C. (2004). Positive development: Realizing the potential of youth. *Annals of the American Academy of Political and Social Science*, 591(6), 202-220. Retrieved from [http://www.jstor.org.ezpdoxy.auckland.ac.nz/stable/4127631?seq=1#page\\_scan\\_tab\\_contents](http://www.jstor.org.ezpdoxy.auckland.ac.nz/stable/4127631?seq=1#page_scan_tab_contents).

- Petrie, K., Penney, D., & Fellows, S. (2014). Health and physical education in Aotearoa New Zealand: An open market and open doors? *Asia-Pacific Journal of Health, Sport and Physical Education*, 5(1), 19-38. doi:10.1080/18377122.2014.867791
- Phelps, E., Balsano, A., Fay, K., Peltz, J., Zimmerman, S., & Lerner, R.M. (2007). Nuances in early adolescent development trajectories of positive and of problematic/risk behaviors: Findings from the 4-H study of positive youth development. *Child and Adolescent Clinics of North America*, 16(2), 473-496. doi:10.1016/j.chc.2006.11.006
- Phelps, E., Zimmerman, S., Warren, A. A., Jelicic, H., von Eye, A., & Lerner, R. M. (2009). The structure and developmental course of positive youth development (PYD) in early adolescence: Implications for theory and practice. *Journal of Applied Developmental Psychology*, 30(5), 571-584. doi:10.1016/j.appdev.2009.06.003
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1-15. doi:10.1016/S0001-6918(99)00050-5
- Punch, K. F. (2009). *Introduction to research methods in education*. London, England: Sage.
- Ramsay, J. O. (1973). The effect of number of categories in rating scales on precision of estimation of scale values. *Psychometrika*, 38(4), 513-533. doi:10.1007/BF02291492
- Rentoul, A. J., & Fraser, B. J. (1979). Conceptualization of enquiry-based or open classroom learning environments. *Journal of Curriculum Studies*, 11(3), 233-245. doi:10.1080/0022027790110306
- Reynolds, C. R., & Paget, K. D. (1983). National normative and reliability data for the revised children's manifest anxiety scale. *School Psychology Review*, 12(3), 324-336. Retrieved from <http://psycnet.apa.org/psycinfo/1984-05662-001>
- Rice, F., Frederickson, N., & Seymour, J. (2011). Assessing pupil concerns about transition to secondary school. *British Journal of Educational Psychology*, 81(2), 244-263. doi:10.1348/000709910x519333
- Rigby, K. (2000). Effects of peer victimization in schools and perceived social support on adolescent wellbeing. *Journal of Adolescence*, 23(1), 57-68. doi:10.1006/jado.1999.0289
- Rigby, K. (2001). Health consequences of bullying and its prevention in schools. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 310-331). New York, NY: Guilford.
- Roberts, W., Hom, A., & Battistich, V. (1995, April). *Assessing students' and teachers' sense of the school as a caring community*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.
- Roehlkepartain, E. C., Benson, P. L., & Sesma, A. (2003). *Signs of progress in putting children first: Developmental assets among youth in St. Louis Park, 1997-2001*. Minneapolis, MN: Search Institute.
- Roeser, R., Eccles, J., & Sameroff, A. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *Elementary School Journal*, 100(5), 443-471. doi:10.1086/499650

- Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Quantitative Methods in Psychology*, 92(3), 726-748. doi:10.1037/0033-2909.92.3.726
- Rossi, P. H., & Freeman, H. E. (1989). *Evaluation: A systematic approach* (4th ed.). Newbury Park, CA: Sage.
- Roth, J. L., & Brooks-Gunn, J. (2003). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science*, 7(2), 94-111. doi:10.1207/S1532480XADS0702\_6
- Roy, K. M., Tubbs, C. Y., & Burton, L. M. (2004). Don't have no time: Daily rhythms and the organization of time for low-income families. *Family Relations*, 53(2), 168-178. doi:10.1111/j.0022-2445.2004.00007.x
- Rumberger, R.W., Larson, K.A., Ream, R.K., & Palardy, G.J. (1999). *The educational consequences of mobility for California students and schools*. Berkeley, CA: University of California.
- Russell, R. (1973). Social health: An attempt to clarify this dimension of well-being. *International Journal of Health Education*, 16, 74-82.
- Saigh, P. A., & Umar, A. M. (1983). The effects of a Good Behaviour Game on disruptive behaviour of Sudanese elementary school students. *Journal of Applied Behavior Analysis*, 16(3), 339-344. doi: 10.1901/jaba.1983.16-339
- Salmond, C., Crampton, P., & Atkinson, J. (2007). *NZDep2006 index of deprivation*. Wellington, New Zealand: Otago University.
- Sameroff, A. (1986). Environmental context of child development. *Journal of Pediatrics*, 109(1), 192-200. doi:10.1016/S0022-3476(86)80604-7
- Sampson, R., & Laub, J. H. (1994). Urban poverty and the family context of delinquency: A new look at structure and process in a classic study. *Child Development*, 65(2), 523-540. doi:10.2307/1131400
- Sanderson, D.R. (2003). Engaging highly transient students. *Education*, 123(3), 600-605. Retrieved from <http://go.galegroup.com.ezproxy.auckland.ac.nz/ps/i.do?action=interpret&ty=as&v=2.1&lm=&u=learn&it=search&s=RELEVANCE&p=AONE&qt=SN~0013-1172~TI~Engaging+highly+transient+students~~VO~123~~SP~600~~IU~3&sw=w&authCount=1>
- Sapsford, R. (2007). *Survey research* (2nd ed.). London, England: Sage.
- Sawhill, I. V., & Karpilow, Q. (2014, July). *How much could we improve children's life chances by intervening early and often?* (Research Brief #54). Retrieved from <http://www.brookings.edu/research/papers/2014/07/improve-child-life-chances-interventions-sawhill>
- Scales, P. (1999). Reducing risks and building developmental assets: Essential actions for promoting adolescent health. *Journal of School Health*, 69(3), 113-119. doi:10.1111/j.1746-1561.1999.tb07219.x
- Scales, P. (2011). Youth developmental assets in global perspective: Results from international adaptations of the developmental asset profile. *Child Indicators Research*, 4(4), 619-645. doi:10.1007/s12187-011-9112-8
- Scales, P. C., Benson, P. L., Leffert, N., & Blyth, D. A. (2000). Contribution of developmental assets to the prediction of thriving among adolescents. *Applied Developmental Science*, 4(1), 27-46. doi:10.1207/S1532480XADS0401\_3
- Scales, P. C., Sesma, A., & Bolstrom, B. (2004). *Coming into their own: How developmental assets promote positive growth in middle childhood*. Minneapolis, MN: Search Institute.

- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54(3), 182-203. doi:10.1037/0003-066X.54.3.182
- Schacter, J. (2001a). *Geographic mobility, March 1999 to March 2000: Current population reports*. Washington, DC: US Census Bureau.
- Schacter, J. (2001b). *Why people move: Exploring the March 2000 current population series - Current population reports*. Washington, DC: US Census Bureau.
- Schaie, K. W., & Strother, C. R. (1968). A cross-sequential study of age changes in cognitive behavior. *Psychological Bulletin*, 70(6), 671-680. doi:10.1037/h0026811
- Schoon, I., Jones, E., Cheng, H., & Maughan, B. (2012). Family hardship, family instability, and cognitive development. *Journal of Epidemiology & Community Health*, 66(8), 716-722. doi:10.1136/jech.2010.121228
- Search Institute. (2011a). *Where the asset framework comes from*. Retrieved from <http://www.search-institute.org/research/assets/background>
- Search Institute. (2011b). *What are developmental assets?* Retrieved from <http://www.search-institute.org/developmental-assets>
- Search Institute. (2011c). *Assets in middle childhood*. Retrieved from <http://www.search-institute.org/developmental-assets-middle-childhood>
- Search Institute. (2012). *Profiles of student life: Attitudes and behaviours*. Minneapolis, MN: The Search Institute.
- Search Institute. (2013). *Developmental assets profile: Technical summary*. Retrieved from <http://www.search-institute.org/sites/default/files/a/DAP-Psychometric-Information.pdf>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Shernoff, D.J. (2010). Engagement in after-school programs as a predictor of social competence and academic performance. *American Journal of Community Psychology*, 45, 325-337. doi: 10.1007/s10464-010-9314-0
- Simmons, R. G., Carlton-Ford, S. L., & Blyth, D. A. (1987). Predicting how a child will cope with the transition to junior high school. In R. M. Lerner & T. T. Foch (Eds.), *Biological/psychosocial interactions in early adolescence* (pp. 375-376). Hillsdale, NJ: Lawrence Erlbaum.
- Simons, R. L., Johnson, C., Conger, R. D., & Lorenz, E O. (1997). Linking community context to quality of parenting: A study of rural families. *Rural Sociology*, 62(2), 207-230. doi:10.1111/j.1549-0831.1997.tb00651.x
- Simpson, G., & Fowler, M. (1994). Geographic mobility and children's emotional/behavioural adjustment and school functioning. *Pediatrics*, 93(2), 303-309. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8121745>
- Sinkinson, M. (2011). Back to the future: reoccurring issues and discourses in health education in New Zealand schools. *Policy Futures in Education*, 9(3), 315-327. doi:10.2304/pfie.2011.9.3.315
- Sinkinson, M., & Burrows, L. (2011). Reframing health education in New Zealand/Aotearoa schools. *Asia-Pacific Journal of Health Sport and Physical Education*, 3(3-4), 53-70. doi:10.1080/18377122.2011.9730359
- Small, S. A., & Rodgers, K. B. (1995). *Teen assessment project (TAP) survey question bank*. Madison, WI: University of Wisconsin-Madison.

- Smith, L. (2004). Changes in student motivation over the final year of high school. *Journal of Educational Enquiry*, 5(2), 64-85. Retrieved from <http://www.cred.unisa.edu.au/jee/Papers/JEEVol5No2/Smith.pdf>
- Smith, J. K., & Hodkinson, P. (2005). Relativism, criteria and politics. In N. Denzin & Y. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.). London, England: Sage.
- Smith, W., & Philpot, R. (2011). The preparation of HPE teachers in New Zealand. *Asia-Pacific Journal of Health, Sport and Physical Education*, 2(3-4), 71-80. doi:10.1080/18377122.2011.9730360
- Sokal, R., & Sneath, P. (1963). *Principles of numerical taxonomy*. San Francisco, CA: W.H. Freeman.
- Solomon, D., Battistich, V., Kim, D., & Watson, M. (1997). Teacher practices associated with students' sense of the classroom as community. *Social Psychology of Education*, 1(3), 235-267. doi:10.1007/BF02339892
- Solomon, D., Battistich, V., Watson, M., Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediated effects of the Child Development Project. *Social Psychology of Education*, 4(3/4), 3-51. doi:10.1023/A:1017107114298
- Sorin, R., & Iloste, R. (2003). *Student mobility: Reasons, consequences and interventions*. Paper presented at the Joint AARE/NZARE conference, University of Auckland, Auckland, New Zealand. Retrieved from <http://www.aare.edu.au/data/publications/2003/sor03064.pdf>
- South, S.J., Haynie, D.L., & Bose, S. (2007). Student mobility and school dropout. *Social Science Research*, 36(1), 68-94. doi:10.1016/j.ssresearch.2005.10.001
- St. John, S. (2012). *Preventing, mitigating or solving child income poverty? The expert advisory group 2012 report*. Auckland, New Zealand: The University of Auckland. Retrieved from <http://igps.victoria.ac.nz/publications/files/b45fba85c9b.pdf>
- Statistics New Zealand. (2005). *Statistical standard for ethnicity*. Wellington, New Zealand: Statistics New Zealand. Retrieved from [http://www.stats.govt.nz/surveys\\_and\\_methods/methods/classifications-and-standards/classification-related-stats-standards/ethnicity.aspx](http://www.stats.govt.nz/surveys_and_methods/methods/classifications-and-standards/classification-related-stats-standards/ethnicity.aspx)
- Statistics New Zealand. (2013). *2013 Census definitions and forms*. Wellington, New Zealand: Statistics New Zealand. Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz)
- Statistics New Zealand. (2014a). *New Zealand in profile 2014: An overview of New Zealand's people, economy, and environment*. Wellington, New Zealand: Statistics New Zealand. Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz)
- Statistics New Zealand. (2014b). *2013 Census quickstats about culture and identity*. Wellington, New Zealand: Statistics New Zealand. Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz)
- Steiger, J.H. (1989). *EzPATH: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.
- Steiger, J. H., & Lind, J. C. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Stevens, J.P. (1979). Comment on Olson: Choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, 86(2), 355-360. doi:10.1037/0033-2909.86.2.355

- Stewart, E. A. (2003). School social bonds, school climate, and school misbehavior: A multilevel analysis. *Justice Quarterly*, 20(3), 575–604. doi:10.1080/07418820300095621
- Strand, S. (2000, September). *Pupil mobility, attainment and progress during key stage 1: A case study in caution*. Paper presented at the British Educational Research Association Annual Conference, University of Cardiff, Wales. Retrieved from <http://www.leeds.ac.uk/educol/documents/00001484.htm>
- Stufflebeam, D. L. (1967). The use and abuse of evaluation in Title III. *Theory into Practice*, 6(3), 126-133. doi:10.1080/00405846709542071
- Stufflebeam, D. L. (1969). Evaluation as enlightenment for decision-making. In H. B. Walcott (Ed.), *Improving educational assessment and an inventory of measures of affective behavior* (pp. 41-73). Washington, DC: Association for Supervision and Curriculum Development and National Education Association.
- Stufflebeam, D. L. (1971, February). *The relevance of the CIPP evaluation model for educational accountability*. Paper presented at the Annual Meeting of the American Association of School Administrators, Atlantic City, NJ.
- Stufflebeam, D. L. (1973). Evaluation as enlightenment for decision-making. In B. R. Worthen, & J. R. Sanders (Eds.), *Educational evaluation: Theory and practice*. Worthington, OH: Charles A. Jones Publishing.
- Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In G. F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation models* (pp. 117-141). Boston, MA: Kluwer-Nijhoff.
- Stufflebeam, D.L. (1994). Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go. *Evaluation Practice*, 15(3), 321-338. doi:10.1016/0886-1633(94)90027-2
- Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation*, 2001(89), 7-98. doi:10.1002/ev.3
- Stufflebeam, D.L. (2003). The CIPP model for evaluation. In T. Kellaghan & D.L. Stufflebeam, *The International handbook of educational evaluation: Part I* (9th ed., pp. 31- 62). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Sudman, S., & Bradburn, N.M. (1974). *Response effects in surveys: A review and synthesis*. Chicago, IL: Aldine.
- Sudman, S., & Bradburn, N.M. (1982). *Asking questions: A practical guide to questionnaire design*. San Francisco, CA: Jossey-Bass.
- Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.
- Tabachnick, B.G., & Fidell, L.S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.
- Takeuchi, D. T., Williams, D. R., & Adair, R. K. (1991). Economic stress in the family and children's emotional and behavioral problems. *Journal of Marriage and the Family*, 53(4), 1031-1041. doi:10.2307/353006
- Te Kete Ipurangi [TKI]. (2014). *Schooling in New Zealand*. Retrieved from <http://www.tki.org.nz/About-this-site/Schooling-in-New-Zealand>
- Theokas, C., Almerigi, J. B., Lerner, R. M., Dowling, E. M., Benson, P. L., Scales, P. C., & von Eye, A. (2005). Conceptualizing and modeling individual and ecological asset components of thriving in early adolescence. *The Journal of Early Adolescence*, 25(1), 113-143. doi:10.1177/0272431604272460

- Toldos, M. P. (2005). Sex and age differences in self-estimated physical, verbal and indirect aggression in Spanish adolescents. *Aggressive Behavior, 31*(1), 13-23. doi:10.1002/ab.20034
- Trickett, E. J., & Moos, R. H. (1973). Social environment of junior high and high school classrooms. *Journal of Educational Psychology, 65*(1), 93-102. doi:10.1037/h0034823
- Trochim, W. M. K., & Donnelly, J. P. (2008). *The research methods knowledge base* (3rd ed.). Mason, OH: Cengage Learning.
- Tsorbatzoudis, H., Travlos, A. K., & Rodafinos, A. (2013). Gender and age differences in self-reported aggression of high school students. *Journal of Interpersonal Violence, 28*(8), 1709-1725. doi:10.1177/08862605112468323
- United Nations. (1989). *Convention on the rights of the child*. Retrieved from <http://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx>
- United Nations International Child Emergency Fund [UNICEF]. (2014). *Children of the Recession: The impact of the economic crisis on child well-being in rich countries*. Retrieved from <http://www.unicef-irc.org/publications/733>
- Utter, J., & Denny, S. (2010). *Youth '07: The health and wellbeing of secondary school students in New Zealand: Differences for young people in neighbourhoods of high and low deprivation*. Auckland, New Zealand: The University of Auckland.
- Vail, K. (1996). Learning on the move. *American School Board Journal, 183*(12), 20-25. Retrieved from <http://eric.ed.gov/?id=EJ540771>
- van Lier, P., Muthen, B., van der Sar, R., & Crinjen, A. (2004). Preventing disruptive behaviour in elementary school children: Impact of a universal classroom-based intervention. *Journal of Consulting Clinical Psychology, 72*(3), 467-478. doi: 10.1037/0022-006X.72.3.467
- VanderVen, K. (2008). *Promoting positive development in early childhood: Building blocks for a successful start*. New York, NY: Springer Science Media.
- Vernon-Feagans, L., Garrett-Peters, P., Willoughby, M., & Mills-Koonce, R. (2011). Chaos, poverty and parenting: Predictors of early language development. *Early Childhood Research Quarterly, 27*(3), 339-351. doi:10.1016/j.ecresq.2011.11.001
- von Collani, G., & Werner, R. (2005). Self-related and motivational constructs as determinants of aggression: An analysis and validation of a German version of the Buss-Perry aggression questionnaire. *Personality and Individual Differences, 38*(7), 1631-1643. doi:10.1016/j.paid.2004.09.027
- Waldegrave, C. (2000, November). Rent policies critical to the offensive against poverty. *The New Zealand Herald*. Retrieved from [http://www.nzherald.co.nz/nz/news/article.cfm?c\\_id=1&objectid=162466](http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=162466)
- Waldegrave, C., & Stuart, S. (1998). *Family dynamics among urban to rural migrants: A study of households moving out of urban areas in New Zealand*. Wellington, New Zealand: Ministry of Agriculture and Fisheries. Retrieved from <http://www.familycentre.org.nz/Publications/PDF%27s/famdyn.pdf>
- Walker, B. J. (2003). The cultivation of student self-efficacy in reading and writing. *Reading & Writing Quarterly, 19*(2), 173-187. doi:10.1080/10573560308217
- Warheit, G., Holzer, C., Bell, R., & Arey, S. (1976). Sex, marital status, and mental health: A reappraisal. *Social Forces, 55*(2), 459-470. doi:10.1093/sf/55.2.459

- Waters, S. K., Lester, L., Wenden, E., & Cross, D. (2012). A theoretically grounded exploration of the social and emotional outcomes of transition to secondary school. *Australian Journal of Guidance and Counselling*, 22(2), 190-205. doi:10.1017/jgc.2012.26
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Weiss, R. S., & Rein, M. (1969). The evaluation of broad-aim programs: A cautionary case and moral. *Annals of the American Academy of Political and Social Science*, 385(1), 133-142. doi:10.1177/000271626938500112
- Weiss, R. S., & Rein, M. (1970). The evaluation of broad-aim programs: Experimental design, its difficulties, and an alternative. *Administrative Science Quarterly*, 15(1), 97-109. doi:10.2307/2391193
- Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38(3-4), 330-336. doi:10.1093/biomet/38.3-4.330
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R.H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues and applications* (pp. 56-75). Thousand Oaks, CA: Sage.
- Williams, B. J., Hay, P.J., & Macdonald, D. (2011). The outsourcing of health, sport and physical educational work: A state of play. *Physical Education and Sport Pedagogy*, 16(4), 399-415. doi:10.1080/17408989.2011.582492
- Wilson, D. (2004). The interface of school climate and school connectedness and relationships with aggression and victimization. *Journal of School Health*, 74(7), 293-299. doi:10.1111/j.1746-1561.2004.tb08286.x
- Wilson, E. B., & Hilferty, M. M. (1931). The distribution of chi-square. *Proceedings of the National Academy of Sciences of the United States of America*, 17(12), 684-688. doi:10.1073/pnas.17.12.684
- Winnheller, S., Hattie, J., & Brown, G. T. L. (2013). Factors influencing early adolescents' mathematics achievement: High-quality teaching rather than relationships. *Learning Environment Research*, 16, 49-69. doi:10.1007/s10984-012-9106-6
- Witt, S. D. (2000). The influence of peers on children's socialization to gender roles. *Early Child Development and Care*, 162(1), 1-7. doi:10.1080/0300443001620101
- Wood, D., Halfon, N., Scarlata, D., Newacheck, P., & Nissim, S. (1993). The impact of family relocation on children's growth and development, school function, and behaviour. *Journal of the American Medical Association*, 270(11), 1334-1338. doi:10.1001/jama.270.11.1334
- Woltman, H., Feldstain, J., Mackay, C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, 8(1), 52-69. No doi.
- World Health Organization [WHO]. (1986). *Ottawa charter for health promotion*. Geneva, Switzerland: World Health Organization.
- World Health Organization [WHO]. (2009). *Milestones in health promotion – Statements from global conferences*. Geneva, Switzerland: World Health Organization.
- Wylie, C. (1999). *Ten years on: How schools view educational reform*. Wellington, New Zealand: New Zealand Council for Educational Research.

- Yeung, A. S., & McInerney, D. M. (2005). Students' school motivation and aspiration over high school years. *Educational Psychology, 25*(5), 537–554. doi:10.1080/01443410500046804
- Yeung, A. S., Lau, S., & Nie, Y. (2011). Primary and secondary students' motivation in learning English: Grade and gender differences. *Contemporary Educational Psychologist, 36*(3), 246-256. doi:10.1016/j.cedpsych.2011.03.001