



Youth Horizons'

Functional Family Therapy Pilot Study

In conjunction with the Christchurch Health and Development study, University of Otago Christchurch

Charles Heywood PhD

November 2014

Youth Horizons Trust
42 Vestey Drive
Mt Wellington
Auckland 1060
New Zealand

Email: Charles.heywood@youthorizons.org.nz

Contents

LIST OF TABLES/ LIST OF FIGURES	iii
EXECUTIVE SUMMARY	v
ACKNOWLEDGEMENTS	v
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: FUNCTIONAL FAMILY THERAPY	3
2.1 Background	3
2.2 Evaluation and meta-analysis	4
2.3 Cost and throughput	7
2.4 Summary	7
CHAPTER 3: THE YOUTH HORIZONS' FUNCTIONAL FAMILY THERAPY PILOT STUDY	9
3.1 Research hypotheses	10
3.2 Research design	10
3.3 Participants	10
3.4 Recruitment and retention	11
3.5 Interviews and questionnaires	13
CHAPTER 4: PARTICIPANT PROFILES	15
4.1 Demographic characteristics	15
4.2 Parental concerns and reasons for undertaking treatment	20
4.3 Parental rationales for agreeing to treatment	22
4.4 Family overview	22
CHAPTER 5: MEASURES AND ANALYSIS	24
5.1 Measures	24
5.2 Attrition	28
5.3 Data analysis	29
CHAPTER 6: YOUNG PERSON CONDUCT PROBLEM AND RELATED OUTCOMES	30
6.1 Outcomes pre to post	30
6.2 Outcomes pre to follow-up	32
6.3 Placement and education outcomes	33
CHAPTER 7: FAMILY FACTORS: PARENTING AND RELATIONSHIPS	35
7.1 Parenting and parent-child relationships	37
7.2 Parent conflict and relationships	38
7.3 Delinquent peer association	38
7.4 Maternal depression	39
7.5 Exploratory multiple regression analyses	40
7.6 Conduct Disorder and Oppositional Defiant Disorder behaviours	40
7.7 Attention-deficit/Hyperactivity Disorder related behaviours	41
CHAPTER 8: OUTCOME COMPARISONS FOR MĀORI AND NON-MĀORI AND BETWEEN GROUPS	43
8.1 Comparison between Māori and Non-Māori: Anti-social behaviour and related outcomes	44
8.2 Comparison between Māori and Non-Māori: Parenting practices	45
8.3 Comparison between Māori and Non-Māori: Parental conflict and relationships	46
8.4 Comparison of outcomes between boys and girls	47

8.5 Comparison of outcomes between older and younger	47
8.6 Comparison of outcomes for greater and lesser severity of behaviour at intake	48
CHAPTER 9: THE INFLUENCE OF MĀORI IDENTITY	49
9.1 Identity and outcomes	49
9.2 Identity and pre-treatment symptom-load	51
CHAPTER 10 PARENT SATISFACTION	53
10.1 Satisfaction with aspects of FFT	53
10.2 Satisfaction for Māori	54
CHAPTER 11 TREATMENT FIDELITY AND OUTCOMES	55
11.1 Measures and procedure	55
11.2 Results	55
CHAPTER 12 SUMMARY AND DISCUSSION	57
12.1 Key findings	57
12.2 Summary of findings related to research hypotheses	58
12.3 Comparison of findings with comparable intervention studies	59
12.4 Discussion of additional outcomes	62
12.5 Moderating factors	65
12.6 Limitations	68
12.7 Future directions	69
APPENDIX 1.	72
REFERENCES	73

List of Tables

Table 1.	Published FFT outcome studies of anti-social behaviour in adolescents	8
Table 2.	Topics covered in the parent pre interview	13
Table 3.	Percentage of parents reporting concern with specific young person behaviours	21
Table 4.	Percentage of parents reporting significant reasons for consenting to FFT treatment	22
Table 5.	Child anti-social behaviours and related outcomes pre to post treatment	30
Table 6.	Child anti-social behaviours and related outcomes pre - treatment to follow-up	32
Table 7.	Placement and education changes before and after FFT	33
Table 8.	Parenting practices and parent child relationship pre to post treatment	35
Table 9	Parent conflict and relationships pre to post treatment	37
Table 10.	Delinquent peer association pre to post treatment	39
Table 11.	Maternal depression severity	39
Table 12..	Multiple regression predicting change in CD/ODD behaviours during FFT intervention	40
Table 13.	Multiple regression predicting change in ADHD behaviours during FFT intervention	41
Table 14.	Child anti-social behaviours and related outcomes: Effect size for Māori and Non-Māori, pre to post	43
Table 15.	Child anti-social behaviour:Effect size by ethnicity, pre to follow-up	44
Table 16.	Parenting practices: Effect size by ethnicity, pre to post	45
Table 17.	Parent conflict and relationships: Effect size by ethnicity	47
Table 18.	Child anti-social behaviour and related outcomes: Effect size for Sole-Māori and Māori-other ethnicities, pre to post	49
Table 19.	Child anti-social behaviour and related outcomes: Effect size for Sole-Māori and Māori-other ethnicities, pre to follow-up.	50
Table 20.	Child anti-social behaviour and related outcomes. Standardised scores (mean 100 SD 10) at pre-test for Sole-Māori, Māori-other and Non-Māori ethnicities	51
Table 21.	Parent satisfaction with aspects of FFT: Percentage of parents' responses	53
Table 22.	Parents of Māori descent: Satisfaction with the cultural sensitivity of FFT	54

List of Figures

Figure 1.	Research design	10
Figure 2.	Participant referral and retention	11
Figure 3.	Young person age at FFT commencement	15
Figure 4.	Young person gender	15
Figure 5.	Primary caregiver ethnicity	15
Figure 6.	Primary caregiver Māori ethnicity	15
Figure 7.	Primary caregiver highest educational qualification	16
Figure 8.	Primary caregiver source of income	16
Figure 9.	Primary caregiver weekly income after tax	17
Figure 10.	Household weekly income after tax	17
Figure 11.	Primary caregiver drinks per week	18
Figure 12.	Percent families with multiple agency involvement	18
Figure 13.	Young Persons by number agency contacts last five years	19
Figure 14.	Comparison with FFT & MST meta-analyses	61

Acknowledgements

I gratefully acknowledge the invaluable assistance of Professor David Fergusson who provided academic advice and guidance for this pilot study. I greatly appreciate the contribution of the 59 young people and their families who participated in this study, and I am grateful for the goodwill of the Auckland Functional Family Therapy team at Youth Horizons who allowed their clients to be interviewed and their outcomes to be analysed. I also acknowledge and thank our funders the Tindall Foundation and Child Youth and Family whose generous support made this project possible. Finally a special thank-you to Mary-Louise Stocker whose editorial support has been considerable and crucial.

Executive Summary

Background to the study

- Conduct problems including conduct disorder and oppositional defiant disorder affect in the region of 5-10% of young people in Aotearoa New Zealand, with males and Māori being at greater risk.
- There is extensive New Zealand and international research to suggest young people with these problems are at increased risk of a wide range of adverse outcomes in adolescence and young adulthood including crime, substance abuse and dependence, motor vehicle collisions, mental health problems, suicidal behaviours, teen pregnancy, and poor physical health.
- Over the last three decades there has been increasing research into effective interventions for the treatment of conduct problems. Functional Family Therapy (FFT) is one programme consistently identified in international studies as an effective intervention.
- FFT is a family therapy intervention programme, grounded in Family Systems Theory, which is used to address a wide range of youth behavioural problems in a range of settings.
- FFT is notable in that it is cost-effective and is able to treat large numbers of families in a shorter period of time than other similar interventions designed to address conduct problems.
- Interventions differ on the extent on which they can be effectively transported from one context or culture to another. In order to reduce conduct problems in New Zealand it is critical that interventions are demonstrably effective for all children, including Māori, and acceptable to all cultures, including Māori.

The Youth Horizons/ University of Otago FFT pilot study

- This pilot study was undertaken by Youth Horizons and the University of Otago to ascertain whether Functional Family Therapy (FFT) in Aotearoa New Zealand was achieving similar results in as those achieved internationally.
- The study was conducted from January 2011-November 2013 with the first FFT team established in New Zealand, at Youth Horizons in Auckland New Zealand.
- A total of 59 families were included in the study. All were referrals from Child Youth and Family, generally for child conduct problems.
- The research design of the pilot was a single group outcome study with repeated measures pre-test, post-test (six months after commencement of FFT) and six month follow up (twelve months after commencement of FFT).
- Each assessment point involved a parent interview, a child questionnaire and a teacher questionnaire. Data from all families who had completed a minimum of one FFT session was analysed.

- The interview and questionnaires covered child anti-social behaviour and drug use as well as parenting and family issues.

Participant profiles

- Forty-five percent of the youth were Māori, the average age was 13 years 7 months at intake, and boys outnumbered girls two to one.
- Over half of the families were sole-parent families and approximately 60% were beneficiaries. About three-quarters of primary caregivers had at most a secondary school education.
- At the commencement of the study forty percent of families reported at least one family member other than the Young Person who had experienced an issue with the Police (21% of families), drugs (14%), depression (33%), anxiety (40%) or suicide attempt (10%) within the last month. Thirty four percent of primary caregivers reported that they had experienced significant periods of low mood or loss of interest in activities within the last two months.
- The three most common parental concerns for the well-being of their child at the beginning of treatment were *missing out on a good job or career*, *behaviour at school*, and *hanging out with other youths who get into trouble*.

Key results

- This pilot study demonstrates that FFT is effective in reducing conduct problems and achieves effects broadly in the medium range. It also confirms that benefits were largely maintained at follow-up. These results are competitive in comparison with outcomes obtained by FFT and similar interventions overseas.
- While benefits for Non-Māori were slightly better than for Māori in some areas, this was not statistically significant.
- All participants including Māori expressed high levels of satisfaction with FFT.
- The findings indicate that that FFT is generally effective for and acceptable to Māori youth and their whānau.

Outcomes for young people

- There was clear evidence that conduct problem and related outcomes improved over the course of the study and, for most measures, continued to do so over the follow-up period. At the end of treatment measures of Conduct Disorder/Oppositional Defiant Disorder (CD/ODD), Attention-Deficit/Hyperactivity Disorder (ADHD), Child Depression and Child Self-reported delinquent behaviour were statistically significant and changed to a degree consistent with a medium effect size. A measure of child alcohol and drug use changed little and Teacher Rated behaviour improved slightly.
- At a six-month follow-up CD/ODD, Teacher rated behaviour and Parent rated delinquent activities remained significant with medium to large effect sizes.

Child alcohol and drug use improved slightly and Child Self-reported delinquent behaviour deteriorated slightly over the follow-up period¹.

- The average effect size for girls was approximately twice that for boys indicating that girls tended to achieve greater improvements in Conduct Problems than boys.
- Older youth (14 and up) obtained effect sizes on average at least twice that of younger (13 and below) on every measure.
- When segregated into older and younger groups it was apparent that child alcohol and drug use remained steady at very low levels amongst younger children but improved modestly amongst older children.
- Youth with above average levels of conduct problems at the commencement of the study obtained larger effects than those with below average levels of problems, across all conduct related measures.
- Living situations changed little over the study period; approximately 75% of children were living with a biological parent with a further 20% living with a caregiver. Very few young persons transitioned to more secure settings such as group homes or Child, Youth & Family secure residences.
- There was no general movement either into or out of mainstream school with the proportion attending school being in the 55-65%.
- Parents reported little change in the degree to which their children associated with other youth who got into trouble, however the young persons reported a significant decrease in the level of anti-social behaviour amongst their friends.

Outcomes for parents/caregivers

A number of parenting and relationship factors were measured pre and post in order to determine what aspects of parent-child interactions were associated with treatment gains.

- Positive parenting (appropriate behaviour management strategies) increased and inconsistent discipline reduced to a significant degree and effect sizes were in the medium range. Corporal punishment also reduced to a degree consistent with a medium to small effect size. Most parenting measures improved although for several the degree of change was in the small range.
- Statistical techniques were used to examine the predictors of change in CD/ODD, and ADHD. Two parenting factors which were consistently associated with improvement were greater parent-child bonding (warmth, attachment) and reductions in hostile discipline practices such as physical discipline and verbal insults.
- There was little change in physical aggression and neglect of the child by the primary caregiver, in part because the rates of these were low to begin with. Primary caregiver verbal aggression to the child (i.e. shouting, insulting) improved to a degree consistent with a small to medium effect size.

¹ ADHD and Child Depression were not assessed at follow-up

- Caregivers reported that their partners made notable gains in terms of reducing verbal and physical aggression and neglect towards the child with effect sizes in the medium range.
- Partner relationships improved with caregivers reporting gains in terms of overall relationship quality (warmth, closeness) and reductions in physical aggression to their partner, both medium effect sizes. Partner violence towards the primary caregiver did not change appreciably.
- Amongst those parents suffering from depression at the commencement of treatment the severity of their depression improved and the rate of depression amongst all caregivers also reduced, but neither change was statistically significant.

Outcomes for Māori

- Non-Māori obtained very slightly larger effect sizes than Māori on average at the end of treatment and at follow-up. Neither difference was statistically significant.
- Non-Māori obtained slightly larger effect sizes for parent-child relationship measures and appropriate parenting and Māori slightly larger effects for (reductions in) inappropriate parenting. In terms of parent-child and parent-partner conflict Māori obtained larger effect sizes and Non-Māori obtained larger effects for caregiver relationship measures. No differences were statistically significant.
- Parents and caregivers who reported being of Sole-Māori descent reported a significantly greater degree of involvement and engagement with Māori culture than those identifying as being of Māori and another ethnicity (Māori-other).
- Whilst treatment progress was not clearly different for Sole-Māori in comparison with Māori-other there were indications that young persons with Sole-Māori parents or caregivers exhibited a slightly lower level of conduct problems in comparison with Māori-other at the beginning of the study.

Satisfaction with FFT

- Participant satisfaction was high with on average 70% of parents reporting being “very much” satisfied with a range of aspects of FFT treatment.
- Over 80% of parents of Māori descent reported being “very much” satisfied with the degree of respect shown to Māori culture, the therapist’s cultural knowledge and the suitability of FFT for Māori whānau.

Recommendations for future research

- Given that the present study lacked a control group, a future study undertaken as a randomised controlled trial would add significantly to the weight and generalizability of the findings.
- Further investigation is required to understand those aspects of FFT associated with different outcomes for Māori, with a view to maximising effectiveness.
- In the current study criminal activity was measured via parent and young-person report. Future studies would benefit from using official crime data given the importance of crime as an outcome and the ability to reliably obtain this data irrespective of the young person's whereabouts.
- Future study designs could include follow-up assessments one or more years post-treatment to ascertain the degree to which treatment gains are maintained in the longer term.

Chapter 1

Introduction

This report describes a pilot study of Functional Family Therapy (FFT) conducted in Auckland, New Zealand by Youth Horizons in conjunction with the University of Otago. Youth Horizons had introduced FFT to New Zealand in 2009 on the grounds that it is a well supported evidence-based intervention for youth with conduct problems. As interventions do not always retain their efficacy when transported from one country or culture to another, empirical verification with a New Zealand population was sought.

1.2 Background to the study

Conduct problems including Conduct Disorder and Oppositional Defiant Disorder afflict approximately 5% to 10% of New Zealand young people, with males and Māori being at greatest risk. There is substantial New Zealand and international research which shows that young people with these problems are at increased risk for a wide range of adverse outcomes in adolescence and young adulthood including: crime, substance abuse, mental health problems, domestic violence and related outcomes.

In accordance with its founding philosophy of evidence-based practice, Youth Horizons has sought to establish internationally developed interventions for youth with conduct problems and their families in New Zealand. Youth Horizons is the first organisation in New Zealand to offer several well known international evidence-based interventions including Multisystemic Therapy (MST), Multidimensional Treatment Fostercare (MTFC), Teaching Family Homes (TFH) and Functional Family Therapy (FFT).

Youth Horizons is aware that interventions proven in one country are not necessarily effective or culturally acceptable in another, and sought the advice of Professor David Fergusson with regards to the planning and implementation of pilot studies of FFT and MTFC, pending a future controlled study.

In consultation with its primary funder, Child Youth & Family, it was agreed that the evaluation of potentially significant interventions, new to New Zealand, was a viable and important project with implications for the future dissemination of services for youth and families.

The following report relates to the Youth Horizons Functional Family Therapy pilot.

Funding for this project was provided by Youth Horizons, Child Youth & Family, the Tindall Foundation and the Lotteries fund. An ethics review was conducted by the Northern X Regional Ethics Committee.

1.3 Report Structure

The report is structured as follows:

Chapter 2 introduces Functional Family Therapy (FFT) and includes a brief meta-analysis of previous studies.

Chapter 3 outlines the research hypotheses to be addressed and describes the study design and procedures.

Chapter 4 outlines the social, economic and demographic characteristics of the participating families.

Chapter 5 describes the measures and statistical procedures used.

Chapter 6 presents the main anti-social behaviour and placement outcomes at the end of treatment and at follow-up.

Chapter 7 presents information relating to parenting and relationship factors which may underlie change in conduct problem outcomes.

Chapter 8 compares behaviour, parenting and relationship outcomes between Māori and Non-Māori, boys and girls, older and younger children, and children with greater or lesser behaviour problems at the outset of treatment.

Chapter 9 evaluates the role of Māori identity in terms of pre-treatment conduct problems and treatment progress.

Chapter 10 presents information relating to the degree to which all parents were satisfied with various aspects of the FFT intervention and the degree to which Māori parents were satisfied with FFT from a cultural perspective.

Chapter 11 investigates the relationship between treatment fidelity and outcomes.

Chapter 12 provides a summary of the findings, an in-depth discussion of the results, the study's limitations and recommendations for future research.

Chapter 2

Functional Family Therapy

This chapter introduces Functional Family Therapy and includes a brief meta-analysis of previous studies.

2.1 Background

Functional Family Therapy (FFT) is the oldest and one of the most widely disseminated of the evidence based interventions for youth conduct problems. Emerging in the late 1960's when there was little hope or expectation that any intervention for conduct problems could be effective FFT has been averred as effective in all significant reviews and is now implemented in 220 sites across five countries and claims to work with more families per year than any other evidence based intervention. Official designations as "evidence based" or "model programme" have been provided by the Centre for Violence Prevention (Blueprints for Violence Prevention), the Office of Juvenile Justice and Delinquency Prevention, the California Evidence Based Clearinghouse for Child Welfare and the Washington State Institute for Public Policy.

FFT marries family systems, behavioural and cognitive-behavioural approaches to intervention in a synthesis based on clinical experience and research-based theory. From Family Systems Theory is derived the concept that the unit of treatment is the family (not just the individual youth) and the family is a system characterised by dynamic relationships between individuals. From the behavioural tradition in psychology comes an emphasis on change in overt behaviour and its immediate causes within the family, and the need to maintain and generalise treatment gains. Allied to this is the use of cognitive-behavioural techniques such as reframing and emotion-management strategies.

FFT treatment is progressed in three phases. The initial phase is termed "engagement and motivation" and is designed to identify and modify intra-family risk factors (e.g. hopelessness, blaming) and strengthen protective factors (e.g. family cohesion) whilst enhancing intervention credibility and family preparedness to change. The second phase is termed "behaviour change" and is focussed on developing individualised strategies for altering cognitive, behavioural and emotional aspects of family functioning (e.g. attributions, parenting skills, managing anger). The final phase is termed "generalisation" and is concerned with maintaining the changes in behaviour achieved in the previous phase, relapse prevention and the use of community resources to facilitate change in contexts beyond the family such as school and sporting groups.

FFT is a short-term intervention of 8-12 sessions on average over a period of 2-4 months. FFT is recognised as very cost effective on a per individual treated basis and for its throughput of families treated.

2.2 Evaluation and meta-analysis

The evidence base for FFT consists of thirteen published studies relating to primary outcomes such as criminal recidivism or substance use (see Table 1)². Of these studies eight were conducted with officially adjudicated juvenile offenders and their families, four with families referred to outpatient substance use treatment programmes and one with families referred to a community based programme for adolescents with behaviour problems. These studies can be divided into a) those conducted by the developers of FFT as opposed to independent researchers and b) efficacy or effectiveness trials where the former relates to studies conducted under ideal conditions and the latter to studies conducted under conditions resembling the typical use of that intervention in the community. For instance efficacy studies will often be conducted by a University, involve highly qualified clinicians and carefully selected participants. In contrast effectiveness studies will tend to be conducted by community agencies, using available staff and with no or limited selection of participants. A well validated evidence based intervention must be demonstrably successful in both efficacy and effectiveness trials.

Considering the seven studies involving juvenile offenders where the primary outcome was official offending data, two were conducted by University based developers of FFT, three were independent University based independent researchers (all efficacy trials), one was a large independent community based trial (an effectiveness trial) and one appears to also be an independent community based trial (effectiveness).³ One of the University based efficacy studies was conducted in Sweden and is thus also a measure of the transportability of FFT to other nations and cultures.

All but one of the University based studies achieved statistically significant reductions in recidivism for the FFT group compared to controls with recidivism ranging from 11% to 60% compared with 47% to 93% in the control groups, amounting to reductions in recidivism of between 29% and 83% relative to control. The odds ratios of control vs. treatment recidivism for these studies ranged from 3.6 to 16 indicating that for every young person receiving FFT who re-offended between 3.6 and 16 young persons receiving the control treatment re-offended. Given that several individuals participating in the control conditions re-offend for every

² For the purposes of this brief meta-analysis a number of studies have been excluded e.g. van der Put et al. 2012 (no treatment as usual control group), Datchi & Sexton 2013 (participants were adult offenders). A list of excluded studies can be provided upon request.

³ This study, Lantz (1982), is briefly described in a reputable secondary source Elliot et al. (2000); the original publication by the Utah Department of Social Services is currently unavailable.

individual receiving FFT this is a *prima facie* reduction in recidivism that is both clinically and socially meaningful. Additionally, following Cohen's the rule of thumb for odds ratios which categorises 1.5 as a small effect, 2.5 as a medium effect and 4.3 as a large effect, this represents a generally large effect on recidivism on statistical grounds.

Outcomes have been robust over time where assessed, with Gordon, Graves & Arbuthnot (1995) following up the Gordon Arbuthnot, Gustafson & McGreen (1988) sample and finding a recidivism rate for FFT of 8.7% compared with 40.9% for the comparison group five years after participating in FFT (odds ratio 7.7; over seven control individuals re-offended for every one who had received FFT).

Of the two community based, independent effectiveness studies Lantz, 1982, (as cited in Elliott, 2000) achieved a recidivism rate of 50% compared to 88% in the control group, resulting in a statistically significant odds ratio of 7.0; seven individuals in the control group re-offended for every individual in the FFT group. However, in the second community based independent effectiveness study, Sexton & Turner (2011) a randomised controlled trial which employed an exceptionally large sample of 917, both treatment and control groups achieved a 22% rate of recidivism equivalent to a non-significant odds ratio of 1.0; for every individual in the control group who re-offended an individual in the FFT group also re-offended. It is noteworthy that the absolute level of recidivism achieved by FFT in this study was bettered in only one other FFT study (Gordon et al 1988; 11% recidivism in the FFT group) and thus the non-significant outcome might be attributed in part to a well performing control treatment.

In a post-hoc analysis Sexton & Turner (2011) determined that those FFT therapists who had a high level of adherence to the FFT treatment model did achieve statistically significant reductions in felony and violent crimes relative to the control group suggesting that FFT delivered by competent therapists is indeed effective. However the point is arguable as it is unknown whether a similar exercise conducted with the control group by considering only the outcomes of those probation officers who were judged more skilful or adherent to the model for usual services would not remove this advantage for the FFT group. It may also be inevitable in real world settings that a proportion of therapists are not highly competent.

When the outcomes of the seven studies involving juvenile offenders are combined statistically using meta-analytic techniques the overall effect for recidivism is statistically significant ($Z = 3.25$, $p < 0.001$) with an overall odds ratio of 4.25, a medium to large effect. The outcomes across the studies were not homogenous ($\text{Chi}^2 = 37.13$, $df = 6$ ($P < 0.00001$)), however they were homogenous (similar) if, and only if, the Sexton & Turner study was excluded. This shows that the recidivism outcome for the latter study was significantly poorer than the outcomes achieved by the remaining studies. As Sexton & Turner (2011) was a community based

effectiveness study this has obvious implications for the dissemination of FFT to community settings⁴.

Only two community-based studies evaluated conduct problem behaviour (as opposed to recidivism); Graham et al. (2013) and White, Frick, Lawing & Bauer (2013); both found effects in medium range for conduct problems in a single group outcome study. When these outcomes were combined the overall effect size was 0.53⁵, a medium effect, and this was statistically significant ($Z = 4.61$, $p < 0.00001$).

Three studies examined the influence of FFT on internalising symptoms such as anxiety and depression; Graham et al. (2013); Rohde, Waldron, Turner, Brody & Jorgenson (2014) and White et al. (2013). Effect sizes ranged from very small to medium. When aggregated using meta-analysis techniques the overall effect size was $d = 0.32$, a small to medium effect, and this was significant ($Z = 2.28$, $p < 0.02$).

The four remaining studies involved substance using adolescents. Friedman (1989) was an independent effectiveness study using community providers. Waldron, Slesnick, Brody, Turner & Peterson (2001) was a University based efficacy study conducted by researchers involved in the development of FFT and Slesnick & Prestopnik (2009) was a University based efficacy study. Rohde et al. (2013) was an efficacy study conducted by a research institute. All were randomised controlled trials but the data for Rohde et al. was calculated pre-post⁶. Friedman (1989) found no significant advantage of FFT compared with alternative treatments including CBT and obtained a small non-significant effect for drug related arrests (Cohen's $d = 0.28$).

Waldron et al (2001) in a comparison of FFT and two alternative treatments obtained favourable results immediately post-treatment but these dissipated to an almost nil effect at the three month follow-up ($d=0.06$). Slesnick & Prestopnik (2009) achieved moderate reductions in alcohol and drug use in comparison with treatment as usual ($d=0.62$). Rohde et al. (2013) obtained a medium effect of 0.46. When the three substance use studies were combined using meta-analytic techniques a medium effect was obtained and this was significant ($d = 0.45$, $Z = 2.57$, $p = 0.01$).

Because of the different measures used, direct comparison of University (efficacy) and community (effectiveness) studies was not possible; for instance none of the University based studies measured conduct problem behaviour. One comparison that can be made is between the two community based studies which assessed recidivism (Lantz, 1982 & Sexton & Turner, 2011) and the five university based studies. Although the data is merely indicative, the mean odds ratio achieved by the latter was 7.3, considerably better than the mean odds ratio for the former studies, 4.0.

⁴ Barnowski (2004) is a large FFT study cited in some reviews; however it appears that the Barnowski paper is an earlier version of Sexton & Turner (2011), and that the two papers refer to the same sample.

⁵ Cohen's rule of thumb for the effect size d are 0.2 = small, 0.50 = medium and 0.8 = large (Cohen, 1988)

⁶ FFT was combined with other treatments, only pre-post sequences of the available data related to FFT only.

Similarly the small number of studies and differing measures does not allow a precise comparison, however the two non-USA based studies Hanson, (1988) and Graham et al. (2013) appear to have obtained outcomes within the range achieved by studies conducted in the United States.

2.3 Cost and throughput

A particular secondary emphasis which distinguishes FFT is a stated concern with achieving ample treatment gains at a lesser cost than other programmes. With FFT treatment teams of between five and six therapists each with caseloads of twelve to fifteen families per team has the potential to treat in excess of 100-150 families a year. This compares very favourably with 50 per Multisystemic Therapy (MST) team and 15 per Multidimensional Treatment Fostercare (MTFC) team and makes FFT an attractive option for social service funders. FFT's cost effectiveness is also very competitive; in a recent cost benefit analysis, FFT for institutionalised juvenile justice offenders had the best net monetary benefit (benefits – costs) and the second best cost-benefit ratio (benefits/cost) of 11 juvenile justice programmes including MST, MTFC and Aggression Replacement Training (ART) (Lee et al. 2012).

2.4 Summary

University based efficacy studies of FFT have reliably demonstrated statistically significant medium to large effects with indications that these outcomes are maintained over time and are internationally transportable to University based settings.

In terms of re-offending FFT typically achieves recidivism rates of less than a quarter of those achieved by alternative treatments or usual services. However the extant literature does not provide evidence for the efficacy or effectiveness of FFT where the primary outcome variable is substance use. Further, two of three community based effectiveness studies, including one very large study, failed to indicate an advantage for FFT over alternative interventions. This may be attributable to lower than optimum treatment adherence in some FFT therapists.

Given the potential value of FFT in terms of low cost, its ability to treat larger numbers of families than most competing interventions and a majority of studies with strong outcomes, FFT is clearly worthy of implementation in New Zealand. However the weaker outcomes relating to substance use and community based programmes also signal the necessity of verifying whether FFT delivers upon its promise in a bi-cultural community setting in New Zealand.

Table 1. Published FFT outcome studies of anti-social behaviour in adolescents

Study	Design, size	Population	Comparison	Outcomes	Odds ratio ¹ or Cohen's d	Provider organisation
Alexander & Parsons (1973)	RCT n = 86	Juvenile offenders	Client centred and eclectic family therapies, no treatment	Recidivism FFT 26% combined alternative treatments and no treatment 55%	OR 3.5	University FFT affiliated
Barton, Alexander, Waldron, Turner & Warburton (1985) Study 3	Non-random control grp n=74	Juvenile offenders	Group homes and other placements	Recidivism FFT 60% alt 93%	OR 9.1	University FFT affiliated
Gordon et al. (1988)	Matched control grp n=54	Juvenile offenders	Probation	Recidivism FFT 11% probation 67%	OR 16	University Independent
Gordon (1995) Study 3	Matched control grp n=52	Juvenile offenders	Probation	Recidivism FFT 33% Probation 64%	OR 3.6	University Independent
Hansson (1998) as cited in Elliott, 2000	RCT n=95	Juvenile offenders	Services as usual	Recidivism FFT 50%, usual services 80%	OR 4.18	University Independent (Sweden)
Lantz (1982) as cited in Elliott, 2000	RCT n=46	Juvenile offenders	Alternative treatment	Recidivism FFT 50%, alternative treatment 88%	OR 7.00	Community Independent
Sexton & Turner (2011)	RCT n=917	Juvenile offenders	Probation services as usual	Recidivism FFT 22%, probation 22%	OR 1.0	Community Independent
White et al. (2013)	Single group n=77	Juvenile offenders	No control group	BASC Rating scale Conduct problems	d = 0.41	Community Independent
White et al. (2013)	Single group n=77	Juvenile offenders	No control group	BASC Rating scale Internalising	d = 0.05	Community Independent
Freidman (1989)	RCT n=121	Substance abusing adolescents	Parent group treatment	Arrests for drug sales (raw data not provided)	d = 0.28 ¹	Community Independent
Graham et al. (2013)	Single group n=98	Adolescents with behaviour problems	No control group	SDQ Conduct problems	d = 0.64	Community Independent (Ireland)
Graham et al. (2013)	Single group n=98	Adolescents with behaviour problems	No control group	SDQ Emotional problems	d = 0.46	Community Independent (Ireland)
Rohde et al. (2014)	Single group n=61	Adolescents with substance use & depression	No control group	Drug use, percent days of use	D = 0.86	Research institute Independent
Rohde et al. (2014)	Single group n=61	Adolescents with substance use & depression	No control group	Depression rating scale	D = 0.46	Research institute Independent
Waldron et al. (2001)	RCT n=114	Substance abusing adolescents	CBT, CBT+FFT, Group therapy	Percentage days marijuana used ² , FFT 40.1 CBT 51.1 Group 41.9	d = 0.06	University Independent
Sleznick & Prestopnik (2009)	RCT n=119	Alcohol abusing adolescents	Ecological family therapy, service as usual	Percentage days alcohol or drug used FFT 13, Service as usual 33 (mean)	d= 0.62	University Independent

¹ The odds relate to the rate of recidivism in the control group vs. FFT for ease of interpretation e.g. odds of 4 = four individuals recidivate for every one in FFT Calculated from F statistic

² Calculated from F statistic

³ Two alternative treatments combined to form control group

Chapter 3

The Youth Horizons' Functional Family Therapy Pilot Study

This chapter outlines the research hypotheses to be addressed and describes the study design and procedures.

3.1 Research hypotheses

The FFT pilot study was intended to supply evidence pertaining to the following questions:

- *Effectiveness.* Does FFT, as delivered by a community agency in New Zealand, lead to beneficial changes from pre to post treatment in key conduct problem related outcomes immediately after the cessation of treatment?
- *Maintenance of gains.* Where beneficial changes are achieved immediately after the cessation of treatment, to what extent are these gains maintained six-months after the cessation of treatment?
- *Equity of outcomes for Māori.* Where beneficial changes are achieved by Non-Māori youth at post-treatment or follow-up, to what extent are these gains shared by Māori?
- *The influence of culture.* For Māori whānau, to what extent is their degree of immersion in Māori culture related to the initial level of conduct problem outcomes at the commencement of treatment and subsequent progress?
- *Characteristics of the population treated.* What are the social, economic, cultural and demographic characteristics of those families receiving the FFT intervention?
- *Mediating and moderating variables.* Which, if any, changes in parent management of youth behaviour from pre to post treatment predict conduct problem outcomes?
- *Treatment fidelity.* To what extent, if any, does therapist to therapist variability in treatment fidelity predict conduct problem outcomes?
- *Client satisfaction – overall.* To what extent are parents satisfied with the FFT intervention?
- *Client satisfaction – Māori.* To what extent do Māori parents find the FFT intervention culturally acceptable?

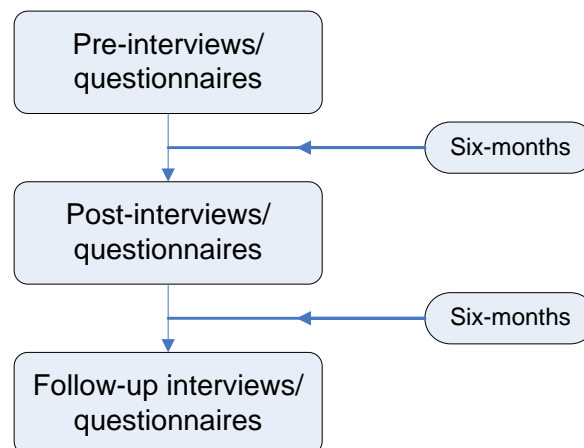
3.2 Research design

The research design of the pilot was a single-group outcome study with pre-test, post-test and a six month follow-up as diagrammed in Figure 1. This design is consistent with the scope of a pilot study and the logistical and practical difficulties associated with a randomised controlled trial.

Each assessment point involved a parent interview, a child questionnaire and a teacher questionnaire (where the child was attending school).

The duration between assessments was set at six months rather than the completion of the intervention to a) accommodate the variable duration of FFT treatment b) to standardise the duration from pre to post to follow-up for analysis purposes.

Figure 1. Research design



3.3 Participants

All referrals to the FFT treatment programme were made by Child Youth and Family (CYF) and all families were required to have at minimum an open file with CYF. Many families also had formal status with CYF such as sections 78 or section 140 care agreements or section 101 or 110 custody agreements.

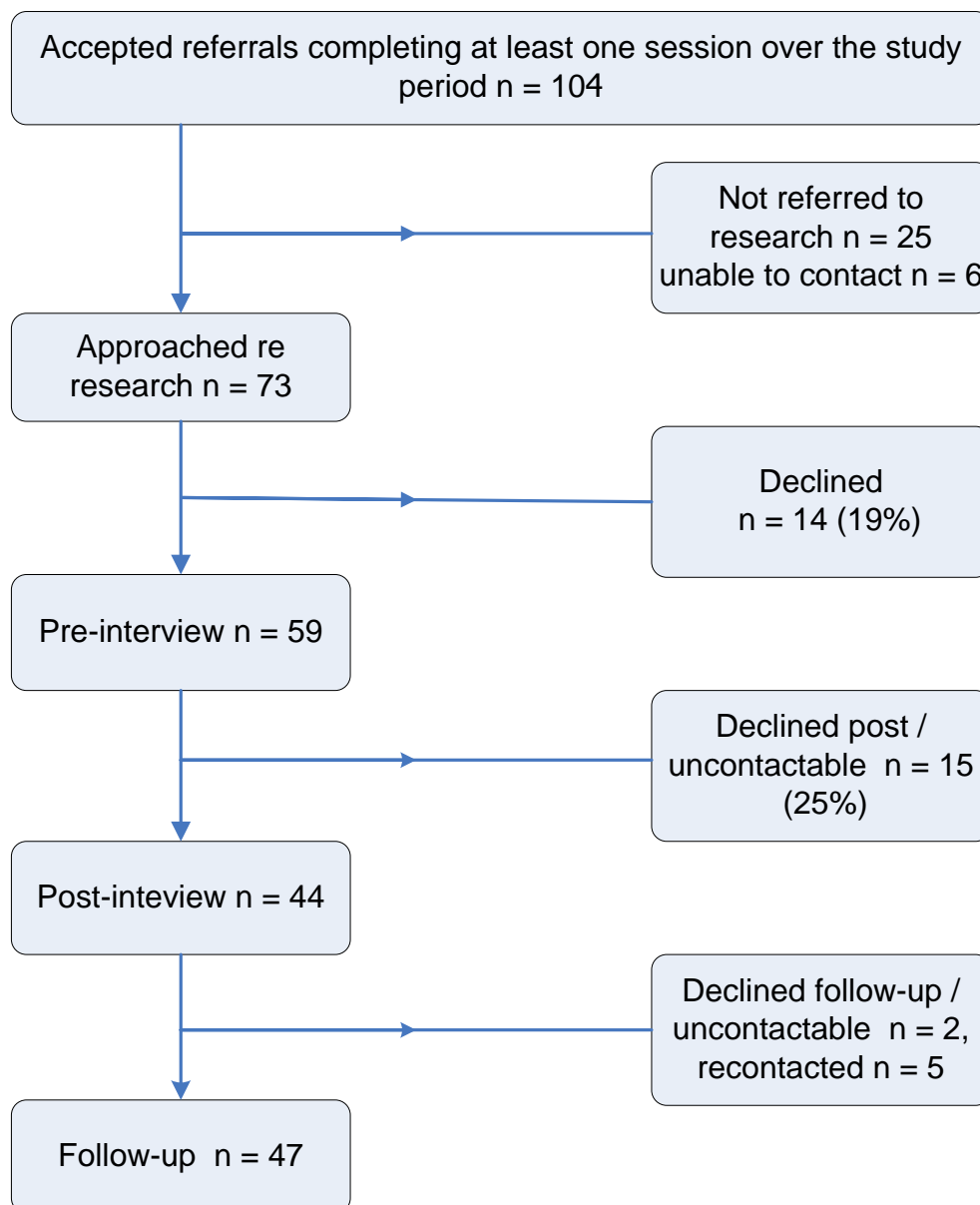
An intention to treat procedure was planned whereby those participants who received FFT as well as those who were intended to receive FFT, but did not, were to be included in the research in order to reduce bias related to treatment drop-outs.

However, it was not considered feasible to define intent to treat as the mere acceptance of a referral by FFT as a considerable number of FFT referrals from CYF were withdrawn over the study period (approximately 50) without being declined by

YHT. Thus using accepted referrals as the criteria for study inclusion could result in only 50% of study participants receiving any FFT.

The revised criteria used for intention to engage in FFT treatment was completion of at least one FFT treatment session. Data for all families who completed at least one FFT treatment session was analysed.

Figure 2. Participant referral and retention



3.4 Recruitment and retention

Overall 59 pre, 44 post and 47 follow-up parent interviews were retained for analysis.

The proposal based on statistical power considerations was to recruit a minimum of 60 families to the study. Written consent was required from both parents and young persons. Only one young person per family was entered into the study in cases where two or more siblings were being treated; where there was a choice the more challenging child by parent nomination was invited to participate in the study.

From February 2011 to October 2012 families were approached as they entered FFT: over that period 104 families commenced FFT. A certain number of families and young persons were not invited to participate as follows, a) following practice in previous studies⁷ 14 families assigned to newly trained therapists with less than three months experience were not invited to participate on the grounds that the risk of poor treatment integrity may be higher in novice therapists, b) 11 families who commenced FFT over the study period were not referred to the research team due to administrative error, delay or press of new referrals. Thus 25 of 104 families were not approached (see Figure 2). Of 79 families to be approached with regards to the research six could not be contacted by the research interviewer in the time available. Of 73 families approached 14 (19%) declined to participate in the research, thus 59 families were pre-interviewed. Over the period from pre-test to post-test, in terms of the parent interview, 15 (25%) families either declined to participate in the post-interview or could not be contacted thus 44 parents were post-interviewed. At follow-up there were two additional families who could not be contacted. However, five families who had missed or refused post interviews were re-contacted at follow-up, amounting to 47 parents in total interviewed at follow-up.

The young person questionnaire was also subject to loss of data with 53 retained at pre-test (59 parent interviews minus 6), 32 at post-test (44 parent interviews minus 12) and 35 at follow-up (47 parent interviews minus 12). Factors known or likely likely to have contributed to this include a) the young person not being present when the parent was interviewed b) the young person's refusal to complete the questionnaire and c) absence of vouchers or other formal incentives.

Forty-three teacher questionnaires were returned at pre-test; 14 young persons were not attending school, one parent did not give consent for the school to be contacted and one questionnaire was not sent in error. Teacher compliance was 100%. Twenty-five teacher questionnaires were returned at post-test; 16 young persons were no longer attending school, two questionnaires were not returned by the school, and two parents did not give consent for school to be contacted. At follow-up 25 teacher questionnaires were retained; 16 young persons were not attending school, two parents did not give consent and three were not able to be secured (child out of area, overseas, no school identified). Whilst the number of teacher questionnaires obtained was much less than the number of parent interviews, by and large this was due to the young person not attending school. School data was not

⁷ E.g. Hogue et al. 2008 only included therapists after four months training

collected where the young person was not attending school at study commencement but returned to school over the course of the study.

Note that all analyses relating to the effects of treatment over time were conducted using pairwise deletion of missing data; that is, for each participant for each measure, only intact pairs (i.e. pre and post) were analysed.

3.5 Interviews and questionnaires

The parent interview

The pre and post parent interview was administered by an experienced research interviewer and took from 60-90 minutes to administer. The follow-up parent interview was designed for administration via telephone and could be completed in ten to fifteen minutes.

Table 2. Topics covered in the parent pre-interview

Section	Topic
A	Family composition
B	Parent education and employment
C	Ethnic identification; for Māori, cultural involvement and engagement
D	School attendance, learning and adjustment
E	Contact with agencies
F to I	Young person behaviour (oppositional defiant disorder, conduct disorder attention-deficit/hyperactivity disorder, depression, delinquent behaviour, drug and alcohol use, age of onset of behavior problems)
J to M	Parenting practices, dealing with and managing misbehavior for primary caregiver and partner
N	Significant life events
O	Maternal/primary caregiver depression
P	Young Person-parent relationship
Q	Young Person's friends
R	Interaction with young person re behaviour
S	Relationship with partner, conflictual interactions with partner
T	Significant issues within immediate family (tobacco, alcohol, drugs, police, anxiety, depression, suicide)
Q	Parent concerns re Young Person, parent expectations of FFT, reasons for referral to FFT

The interview consisted of sections written for the interview as well as sections sourced from other instruments. The topic domains are outlined in Table 1. Detailed

descriptions of the measures in each section are given in Chapter 5. The post-interview was largely the same except a), section C ethnic identification was omitted and b), section Q was changed to family satisfaction with FFT and reasons for leaving FFT where this occurred. The follow-up parent interview was designed to be brief and included only sections F, H, I and added new sections relating to the young person's current living situation, education and employment.

Young Person and teacher questionnaires

The young person questionnaire was identical to the parent interview for Section I Delinquent Activities but was framed as a self-report. In addition, at pre-test and post-test, young persons were asked if their friends engaged in these delinquent activities.

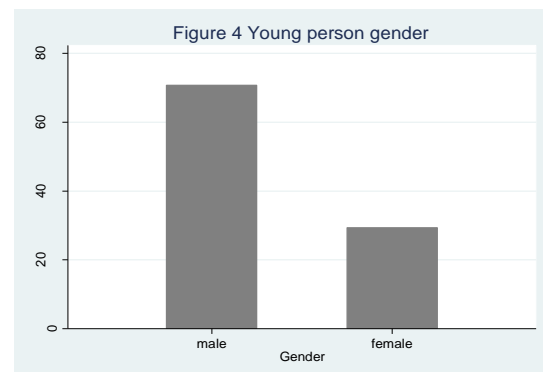
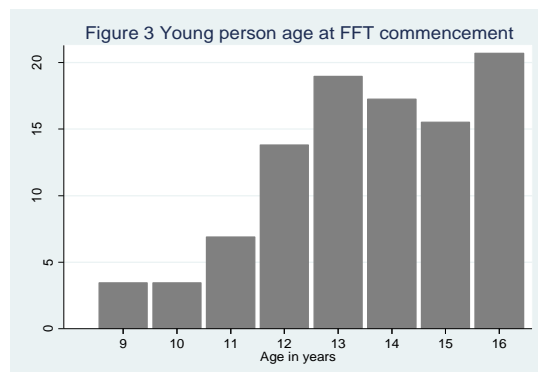
The teacher questionnaire was identical to the parent interview for Section F Oppositional Defiant Disorder and Conduct Disorder subsections. Items relating to school attendance and attitude to school were added to the teacher questionnaire.

Chapter 4

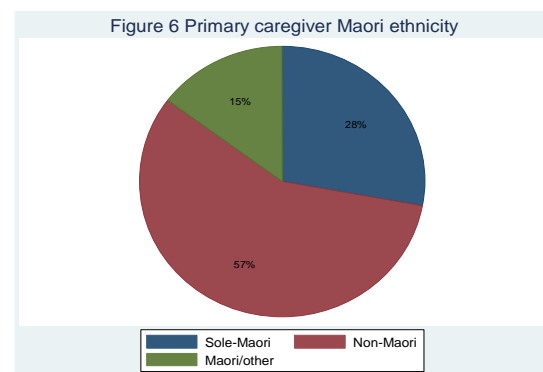
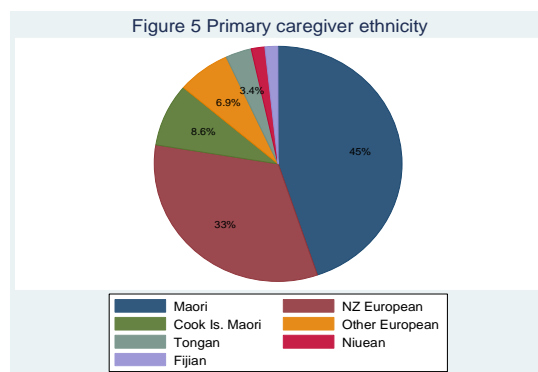
Participant profiles

This chapter outlines the social, economic and demographic characteristics of the participating families.

4.1 Demographic characteristics



The mean age of young persons entering the study was 13 years 7 months with a range from 9 to 16 years. As can be seen in Figure 3 the majority of young persons were within the 12 to 16 year range. The sample was predominantly male and boys outnumbered girls two to one (70% vs. 30% of the sample).



Ethnicity was based on parent or primary caregiver self-report. Interviewees were able to nominate more than one ethnicity. Thus some individuals identified solely as Māori (Sole-Māori 29%, see Figure 6) and others as Māori and another ethnicity (Māori/other 15%). Sole-Māori and Māori-other ethnicities are combined in Figure 5 and it can be seen that Māori are the largest ethnic group represented in this study (45%) followed by New Zealand European (33%). "Other European" refers to persons identifying as English, Dutch, German or other European nationalities (7%).

The remaining ethnicities include Cook Island Māori (10%), Tongan (3 %), Niuean (2%) and Fijian (2%).

Of note is the complete absence of any families of Indian or Chinese descent despite persons of Asian origin making up 19% of the population in the Auckland region in the 2006 census. Despite Māori constituting 45% of this sample, the overall Māori population in Auckland sits at a far lower rate of 11%. Pasifika representation mirrored the population rate; 14% of the Auckland population and 17% of the current study.

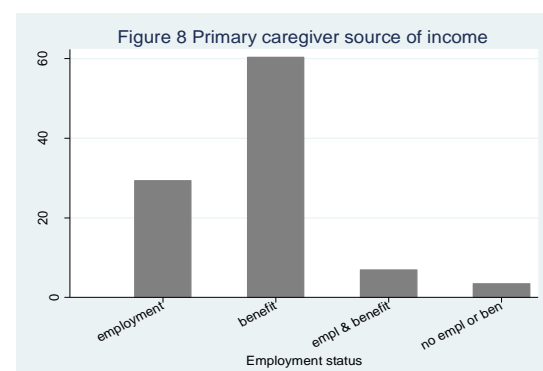
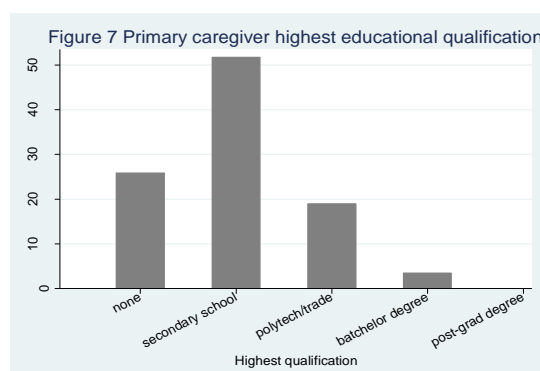
Parents were also asked if they were of Māori descent given that a person may identify as Non-Māori but have Māori ancestry; 48% of interviewees affirmed this, closely following the proportion of those identifying as Māori or part-Māori. Sole-Maori or Māori-other.

Family composition

Thirty-one percent of families included two parents or caregivers and sixty-nine percent were sole-parent families. Seventy-six percent of primary caregivers were biological parents, twenty-one percent were adoptive or step-parents. Of the thirty-three percent of two-parent families, fifty-two percent were adoptive or step-parents and the remainder biological parents. The great majority of households (eighty-percent) included one or more siblings of the client young person; the median number of siblings was two. Of siblings, sixty-two percent were full biological siblings and thirty-eight percent half, adopted, step or foster siblings.

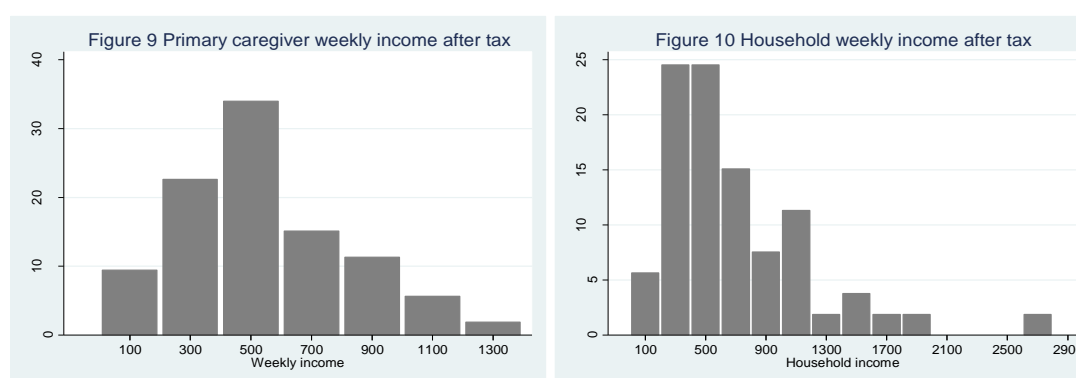
Education, income & employment

More than three-quarters (78%) of the primary caregivers of those families receiving FFT either had at most a secondary school qualification and approximately a quarter had no qualification at all (see Figure 7). The most prevalent source of income was a social welfare benefit (62%), and approximately one third (38%) of caregivers were working in paid employment. Of those receiving benefits three-quarters (75%) were on the Domestic Purposes Benefit with smaller numbers receiving Invalids, Sickness or other benefits.



The primary caregivers' median total weekly income from all sources was \$500, (range \$0 to \$1400, mean \$525) equivalent to an annual income of \$26,000 (see Figure 9); this compares with the national median weekly income of \$28500 per annum or \$548 per week in 2013. (Note that the census data will have included significant numbers of persons with no employment such as students and retirees).

Household income was more variable, however the median weekly income was also \$500 (range \$100 to \$2700, mean \$622) compared to the national median weekly household income of \$1525 in 2012 (see Figure 10).⁸ Thus primary caregiver incomes were slightly below the national median while household incomes were considerably less, possibly due to the high proportion of sole-parent households and the consequent lack of a second income.



New Zealand Socio-economic Index (NZSEI) 2006 scores were estimated for each household based on occupation, or education (if no occupation).⁹ As so many parents were beneficiaries education was the more commonly used criterion. The median score was 41 which is within in the 4th highest of six socio economic groups as used in the Elley-Irving socio-economic scale. The average participant household in this study was of higher socio-economic status than approximately 40% of New Zealand households based on the occupation or education of the primary breadwinner. Thus the socio-economic status of this sample was not markedly low but this may reflect the NZSEI guidelines for assigning scores on the basis of education.

Family stressful life events

Families in at-risk social groups are known to experience an elevated level of stressful transitions and events such as moving house, becoming unemployed, accidents, illness and financial difficulties. The families commencing FFT reported a median of four life events in the previous year (range nil to fifteen) which they rated as either "upsetting" or "very upsetting". Over one-third (37%) experienced the

⁸ New Zealand Statistics Income Survey 2013

⁹ New Zealand Socio-economic Index 2006, Statistics New Zealand (2013)

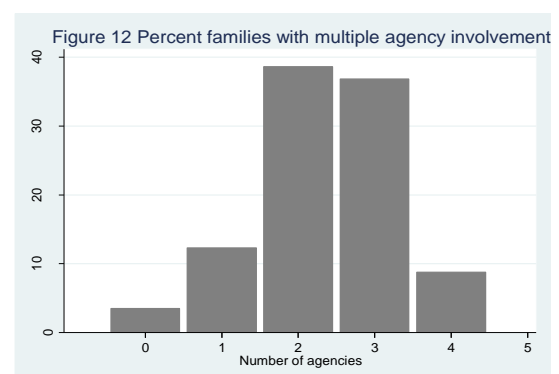
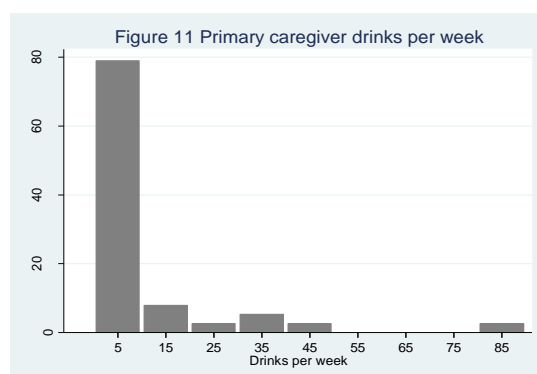
death of a close friend or family member. One half experienced serious illness or injury of themselves or family member. A similar proportion (46%) reported serious arguments with family members or spouse. Over one-third (37%) experienced serious financial problems. Smaller but significant proportions of respondents reported becoming unemployed (9%) and being assaulted by a partner (5%).

Family social issues

In terms of issues within the immediate family 40% of families reported at least one family member other than the young person who had experienced an issue within the last month with the Police (21% of families), drugs (14%), depression (33%), anxiety (40%) or suicide attempt (10%). In terms of depression 34% of primary caregivers reported that they had experienced significant periods of low mood or loss of interest in activities within the last two months; 27% reported nine or more symptoms of a depressive episode during these periods and 24% reported that this interfered with at least one activity of daily living (employment, relationships) "a great deal". These rates are conspicuously in excess of those expected in the general population. Of female primary caregivers 55% smoked cigarettes as well as 19% of male caregivers, 16% of young persons' brothers and 10% of sisters.

Family alcohol use

About one quarter (26%) of primary caregivers reported consuming alcohol in a typical week and 15% reported modest consuming alcohol daily. However, it can be seen in Figure 11 that a minority of caregivers report significant alcohol consumption. In 19% of families the primary caregiver and in 12% the caregiver's partner were reported to have consumed more than six drinks in one session within the last month.

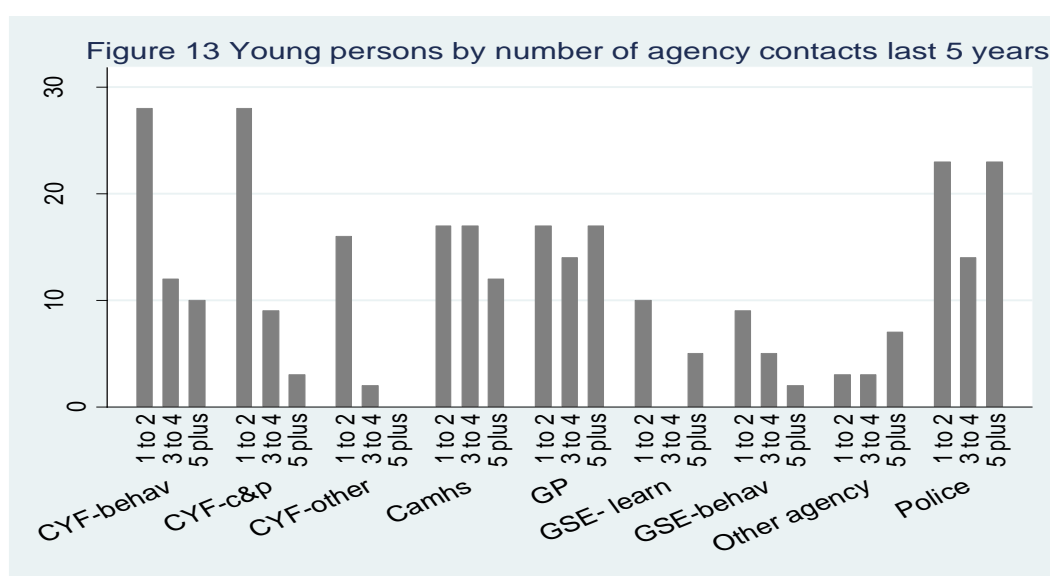


Multiple agency contact

Figure 12 shows the level of multiple agency contact: the median number of agencies involved with the families in this study is two, and almost half of families have contact with three or more agencies¹⁰. This suggests that the young persons and families have pervasive difficulties and are presenting to Group Special Education (GSE), Child & Adolescent Mental Health Services (CAMHS), the police and other organisations prior to being referred to Youth Horizons. The most common origin of the most recent referral to CYF was Family Group Conference (37% of referrals) followed by the Police (24%), school (19%), parents (19%) and CAMHS (17%)¹¹.

Contact with agencies

Figure 13 depicts the number of agency contacts experienced by families; it can be seen that over a quarter of families had received one-to-two contacts with CYF with regards to child behaviour or care and protection issues, approximately one-half (46%) had contact with CAMHS and approximately one-quarter (23%) had had contact with GSE over child behaviour or learning issues. Approximately half (48%) reported contact with their general practitioner (GP) over this period. Of particular note is the level of contact with the police, with over half having at least one police contact and almost a quarter (23%) with five or more contacts.



¹⁰ GP contact is excluded as it is a positive index rather than a measure of family adversity.

¹¹ Some caregivers indicated more than one source of the most recent referral

School

Sixty-one percent of young persons were attending school at the beginning of this study; one young person was attending correspondence school, and the remainder were not attending school. Of those attending school, parents reported that 23% were experiencing a "serious" problem with reading and an additional 14% a mild problem with reading. Similarly, 20% and 26% respectively were having "serious" problems with English and Mathematics and 17% and 20% were having "mild" problems. Overall 57% of those attending school reported some degree of academic problem. Of those not currently attending school 74% were reported to have had some degree of academic problem when they last attended school. Of those attending school 46% were reported to have a "serious" problem with behaviour and/or "serious" difficulty getting along with his/her teacher. Forty-eight percent were reported to have engaged in some degree of truancy or school refusal and in eight cases, or 23% of those attending school, the truancy or school refusal involved 11 or more absences over the last six months.

4.2 Parental concerns and reasons for undertaking treatment

Parents and primary caregivers were asked to what extent specific risks associated with adolescent conduct problems concerned them. Parents averred on average eight of sixteen risk domains as being of concern "a great deal"; this is suggestive of a very high level of stated parental or caregiver concern with the behaviour of the child in their care. The three most commonly indicated concerns were missing out on a good job or career, behaviour at school and getting expelled from school. These point towards a concern for the young person's future prospects as an adult, which is dependent on education.

The risks of association with delinquent youth, well known in the academic literature, are acknowledged by parents with 58% having a "great deal" of concern as to the influence of other youth on their child. Although about half of parents and caregivers have a "great deal" of concern regarding the behaviour of their child at home, 30% are "somewhat" concerned (the largest percentage in this category); this suggests that severe behaviour problems in the home are not universal and in some cases that issues in the home as witnessed by parents may be mild.

Going to prison, getting into crime and getting into trouble with the police are concerns, but not predominant amongst the range of issues parents considered. Given that the mean age of young persons commencing FFT was 13 years, crime and prison may seem distant prospects compared to getting expelled from school or getting into trouble with other youth. Getting involved in gangs is not frequently cited; organised gangs are possibly more of a risk for an older cohort.

Table 3. Percentage of parents reporting concern with specific young person behaviours

Concern	not concerned	somewhat concerned	concerned a great deal
Missing out on a good job/career	21%	16%	63%
Behaviour at school	26%	12%	61%
Getting expelled from school	32%	9%	60%
Hanging out with other youth who get in trouble	28%	14%	58%
Getting into fights & getting hurt	28%	14%	58%
Behaviour at home	19%	30%	51%
Getting pregnant/getting a girl pregnant	33%	18%	49%
Having sex too young	33%	18%	49%
Getting in to drugs	33%	21%	46%
Going to prison	49%	7%	44%
Getting hurt due to risky behaviour e.g. getting in car with drunk driver	39%	18%	44%
Getting into crime	39%	19%	42%
Getting into trouble with the Police	43%	16%	42%
Behaviour in the community	32%	28%	40%
Getting involved with gangs	58%	7%	35%
Violent behaviour to others	46%	28%	26%

4.3 Parental rationales for agreeing to treatment

The great majority of parents and caregivers were motivated to consent to FFT treatment by a desire to receive help both for the young person and in order to better manage the young person. There was also a significant desire for assistance with more general family problems and this is apt for a treatment which is self-described as a form of family therapy. Over half of parents report receiving a directive or instruction to undertake FFT; clearly the majority are also averring a need for help with the young person's behaviour but there is a perhaps an undercurrent of actual or perceived compulsion. Advice from Kaumātua or professionals was not commonly cited.

Table 4. Percentage of parents reporting significant reasons for consenting to FFT treatment

Reason	somewhat significant	significant a great deal
Help for Young Person re his/her behaviour	7%	91%
Help for family to cope with Young Person's behaviour	9%	89%
Help for family problems in general	9%	86%
A directive from CYF, an FGC or the courts	16%	61%
FFT or Youth Horizons helped other families you know	4%	11%
Advice from professionals that the Young Person would benefit from FFT	5%	23%
Advice from Kaumatua or elder that the Young Person would benefit from FFT	2%	9%

4.4 Family overview

The young person participants were on average 13 years 7 months of age at the commencement of FFT and boys outnumbered girls two to one. Māori were the largest single ethnicity represented in the study and comprised 45% of participants.

The participating families exhibited multiple indices of social disadvantage. Approximately two-thirds of families were sole parent families. In over half of families the main source of income was a social welfare benefit. Levels of education amongst parents were low with three-quarters having at most a high school education. Personal income levels were slightly below the national median but household income was approximately one third of the national median.

Families reported a median of four stressful life events over the previous year as well as elevated levels of financial problems, illness and conflict within the immediate family. Alcohol consumption was generally modest but a minority of parents reported binge drinking within the previous month. Over 50% of families had had contact with two or more government agencies and a minority had had multiple contacts with one or more agencies. Approximately one third of parents had a current or recent live-in partner and almost two-thirds of these reported some degree of verbal abuse from a partner. Six percent reported some degree of physical abuse from a partner.

Over one third of young persons were not attending school at the commencement of the study. Of those attending school just under half were reported to have serious academic and/or behaviour problems.

Chapter 5

Measures and analysis

This chapter describes the measures and statistical procedures used.

5.1 Measures

5.1.1 Conduct problem and associated measures

Seven conduct problem related measures were collected pre-test, post-test and at six month follow-up.

1. *Parent rated Oppositional Defiant Disorder and Conduct Disorder CD/ODD-P*. This measure was based on a sum of 23 items derived from the DSM-IV criteria for Oppositional Defiant Disorder and Conduct Disorder. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.88¹².
2. *Parent rated child tobacco, alcohol and drug use CAD-P*. This measure was based on the number of days per year of use across eight classes of substance, summed across substances. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.74.
3. *Parent rated child delinquent activities DEL-P*. This measure was based upon the estimated frequency over the previous six-months of 31 anti-social behaviours such as hitting a parent, being suspended from school and using force or threats to obtain money, summed across behaviours. The measure was derived from a very similar scale used in the Christchurch Health and Development Study. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.71.
4. *Young person self-report of delinquent activities DEL-YP*. This measure was almost identical to the DEL-P except that it was a self-report from the young person. It was based on the estimated frequency of 35 anti-social behaviours over the previous six months and the scale's reliability within the present sample as assessed by Cronbach's alpha was 0.69.
5. *Teacher rated Oppositional Defiant Disorder and Conduct Disorder related behaviours CD/ODD-T*. This measure was based on a sum of 23 items derived from the DSM-IV criteria for Oppositional Defiant Disorder and Conduct Disorder and differed only slightly from the CD/ODD-P measure. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.87.

¹² Reliabilities of 0.8 or more are considered good and 0.7 or more adequate in social science research.

6. *Parent rated Attention-deficit/Hyperactivity Disorder related behaviours ADHD.* This measure was based on the sum of 18 items derived from the DSM-IV criteria for ADHD. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.88.
7. *Parent rated child depression, DEP.* This scale was based on the DSM-IV criteria for major depressive episode augmented with items of particular relevance to children and adolescents. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.91.

5.12 Parenting measures

The study used 19 parenting and related measures across five categories.

Scales relating to Parental involvement, Positive parenting, Poor supervision, Inconsistent discipline and Corporal punishment were derived from the Alabama Parenting Questionnaire (Shelton, Frick & Wootton, 1996).

1. *Parental involvement.* This scale was based on ten items relating to the degree to which parents engaged in positive interactions with their children such as playing games, helping with homework and attending activities together. Dadds, Majeaun & Fraser (2003) recorded an internal consistency of 0.75 and a two-week test retest reliability of 0.87, both acceptable values.
2. *Positive parenting.* This scale was based on six items reflecting the parent's use of positive discipline practices such as praise and rewards. Dadds et al. (2003) recorded an internal consistency of 0.77 and a two-week test retest reliability of 0.85, both acceptable values.
3. *Poor supervision.* This scale was based on ten items assessing the degree to which the parent failed to adequately monitor their child's behaviour such as allowing the child to go out with friends they do not know and having no set time to return home. Dadds et al. (2003) recorded a modest internal consistency of 0.59 and an acceptable two-week test retest reliability of 0.84.
4. *Inconsistent discipline.* This scale was based on six items relating to the parent's sporadic or unreliable use of discipline such as threatening a punishment which is not delivered or varying a punishment depending on the parent's mood. Dadds et al. (2003) recorded a moderate internal consistency of 0.73 and a good two-week test retest reliability of 0.88.
5. *Corporal punishment.* This scale was based on three items relating to the use of physical punishment including slapping, smacking and hitting the child with a belt or cane. Dadds et al. (2003) recorded a weak internal consistency of 0.55 and a very good two-week test retest reliability of 0.90.

Scales relating to Lax discipline, Over-reactive discipline, Hostile discipline and Appropriate discipline were derived from the Parenting Scale (Arnold, O'Leary, Wolff & Acker, 1993).

6. *Lax discipline.* This scale was based on five items relating to unassertive discipline such as not following through on requests and not responding straight away when the child does something inappropriate. Rhoades & O'Leary (2007) reported moderate internal consistency of 0.74 for the mother's report.
7. *Over-reactive discipline.* This scale was based on five items relating to the parent's verbose or excessive response to child misbehaviour such as lecturing the child or raising his/her voice. Rhoades & O'Leary (2007) reported a moderate internal consistency of 0.67 for the mother's report.
8. *Hostile discipline.* This scale was based on three items relating to verbally or physically aggressive responses to child misbehaviour such as smacking, shouting or insulting the child. Rhoades & O'Leary (2007) reported a moderate internal consistency of 0.78 for the mother's report.¹³

5.13 Measures relating to conflict and harmful parenting

Scales relating to parental verbal and physical aggression towards the child, as well as neglect and appropriate non-violent discipline were derived from the Revised Conflict Tactics Scale (Straus et al., 1998). The primary caregiver was asked initially to report on their own behaviour and then to report on the behaviour of their live-in partner or spouse.

9. *Non-violent discipline.* This scale was based on four items relating to appropriate alternatives to physical punishment such as time out or loss of privileges. Straus, Hamby, Finkelhor, Moore & Runyan (1998) cite an adequate internal consistency of 0.70 for this scale.
10. *Verbal aggression.* This scale was based on five items relating to verbal and symbolic acts on the part of the parent with the intent of causing psychological pain or fear such as threatening, yelling or insulting the child. Straus et al. (1998) cite a modest internal consistency of 0.60 for this scale.
11. *Physical aggression.* This scale was based on eleven items relating to a range of parental acts of physical aggression such as punching, slapping or striking the child with an object. Straus et al. (1998) cite a low to moderate internal consistency of 0.55 for this scale.
12. *Neglect.* This scale was based on four items relating to parental failure to engage in tasks necessary for the well-being of the child such as failing to ensure adequate

¹³ Internal consistency is inversely related to the length of the scale and assessments with a small number of items only will be inclined to achieve low consistency, all else being equal. Thus low internal consistencies for short scales may be due to length rather than poor quality items per se.

food was available or failing to take the child to the doctor when required. Straus et al. (1998) cite a low internal consistency of 0.22 for this scale.

5.14 Measures relating to the quality of the parents' relationship

Items relating to inter-partner attachment were derived from the Christchurch Health and Development Study 21 year interview (CHDS 21) which had adapted items from the Partner Attachment Scale (Braiker & Kelly, 1979). Questions relating to inter-partner violence were derived from the Conflict Tactics Scale (Strauss, 1979).

13. *Partner relationship quality.* This scale was based on twenty-six items reflecting the emotional bond, attachment or closeness the primary caregiver feels for his/her partner. The scale also touches upon relationship difficulties such as arguments. Internal consistency in the present sample was high, 0.94.
14. *Partner violence.* This scale includes twenty items reflecting the extent and severity of violence directed at the primary caregiver by their live-in partner. A range of actions were surveyed such as hitting, pushing, punching and using a weapon. Internal consistency in the present sample was good, 0.90.
15. *Violence to partner.* This scale includes 13 items, many of them taken from the previous scale, indicative of violence directed at a partner by the primary caregiver. Internal consistency for this scale in the present sample was poor, 0.35, however examination of the raw data showed that on 87% of occasions the scale total was nil.

5.15 Measures relating to parent-child bonding

Items relating to the emotional bond, or lack thereof, between parent and child were derived from the Christchurch Health and Development Study 15 year interview (CHDS 15) questionnaire.

16. *Parent-child bonding.* This scale included fourteen items relating to trust, conflict and closeness occurring between parent and child. Internal consistency in the present sample was 0.85.

5.16 Measures relating to deviant peer association

Items relating to the young persons were derived from the CHDS 15 year questionnaire.

17. *Deviant peer association.* This scale included five parent report items relating to the degree to which the young person's friends engaged in anti-social activities and encouraged the young person to do likewise. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.94.

18. *Friends' delinquent activities*. This scale is identical in content to the DEL-YP measure described above except that the young-person indicates if a best friend, or any other friend, has engaged in any of the 31 anti-social behaviours surveyed. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.96.

5.17 Measures relating to maternal depression

Items relating to maternal depression were derived from the CHDS 15 year questionnaire.

19. *Maternal depression*. This scale consisted of 30 items relating to the presence of effects of depressive symptoms as reported by the primary caregiver. The scale's reliability within the present sample as assessed by Cronbach's alpha was 0.84.

5.2 Attrition

Of the 59 parents who completed the pre-test interview 17 (29%) missed one or other of the post-test or follow-up interviews. In 12 of the 17 cases (70%) the family did not complete treatment and were no longer compliant with or available for either research or treatment. One pre-interview was lost in a house burglary and was not able to be repeated thus all analyses of participants at the commencement of the study were based on 58 interviews.

Of those families who completed treatment and missed any of the pre, post or follow-up assessments the majority were missing the follow-up assessment. In order to evaluate any potential mortality bias due to any systematic difference between those participants with missing data (the majority being treatment drop-outs) and those with complete data (the majority being treatment completers) scores at pre-test were compared for each measure between those participants who completed all assessments and those missing at least one assessment. Pre-test average scores for those missing data vs. not-missing data were very similar for parent rated CD/ODD, ADHD and Depression ratings.^[1]

Child alcohol and drug use, parent rated delinquent activities and teacher-rated CD/ODD behaviours all evidenced larger (more severe) scores for those with missing data, however the differences were not statistically significant. Young person self-reported delinquent activities displayed more severe scores at pre-test for those with complete data although this was not significant. Overall there was some evidence, albeit inconsistent, of greater symptom severity at the beginning of treatment amongst those participants who subsequently dropped-out or otherwise missed assessments.

^[1] Numerical data and t-tests for this section are available upon request.

5.3 Data analysis

Pairwise deletion of missing values was used such that all comparisons were of groups of equal numbers. Each comparison was tested for statistical significance using the t statistic for paired data. Effect sizes were estimated using Cohen's d for unpaired data in order to retain comparability with other studies using this statistic.

Multivariate analysis of variance (MANOVA) was used to statistically combine data from groups of related measures and test for overall significance. The MANOVA statistic of interest was the interaction effect of measurement by time which reflects the degree of change in the measures from pre to post or pre to follow-up. Multivariate eta squared was used as an estimate of the effect size across each group of measures. For this analysis multiple imputation was used to estimate missing data points and allow the inclusion of data from all 59 participants. Teacher data was not included in these analyses as approximately 40% of young persons did not attend school and it would be inappropriate to impute teacher ratings as if they did.

The validity of the MANOVA analyses depended on two assumptions; firstly that the missing data was missing at random and secondly that skew present in certain measures was not unduly influencing outcomes. Given the previous section there is evidence of some degree of bias in some measures suggesting that missing data for these measures may not be missing at random. However, minor violations of the missing at random requirement may have minimal influence and the imputation procedure greatly adds to the statistical power of the analysis. In order to evaluate any potential influence due to data not missing at random each MANOVA was repeated with a) no imputed data and b) with a dataset in which truncation had been used to minimise and eliminate skew. Where the conclusions of each analysis are concordant it can be reasonably concluded that issues of skew and data possibly not missing at random do not have a significant influence on the outcome.

Chapter 6

Young person conduct problem and related outcomes

This chapter presents the main anti-social behaviour and placement outcomes at the end of treatment and at follow-up.

6.1 Outcomes pre to post

Table 5 compares mean scores on seven measures related to childhood conduct problems. Each comparison is tested for statistical significance using the t statistic and the size of effect is estimated using Cohen's d.

Table 5. Child anti-social behaviour and related outcomes pre to post treatment

Measure	Pre mean (SD)	Post mean (SD)	(df) t	p	d (95% CI)
CD/ODD ¹ (parent)	102.6 (9.5)	97.3 (9.9)	(42) 3.85	0.0004**	0.55 (0.11 to 0.98)
Child alcohol/drug use ¹	100.1 (8.8)	99.9 (11.2)	(42) 0.12	0.90	0.02 (-0.40 to 0.44)
Child delinquent activities (parent) ¹	100.6 (9.4)	99.4 (10.6)	(42) 0.70	0.49	0.12 (-0.30 to 0.55)
Child delinquent activities (self report) ²	102.8 (10.8)	97.2 (8.4)	(28) 2.46	0.02*	0.58 (0.06 to 1.1)
CD/ODD (teacher) ³	101.05 (8.7)	98.9 (11.2)	(24) 0.88	0.38	0.21 (-0.34 to 0.76)
ADHD ¹	102.2 (9.6)	97.8 (10.0)	(42) 3.68	0.0007**	0.45 (0.02 to 0.87)
DEP ¹	102.3 (10.2)	97.7 (9.3)	(42) 3.33	0.002**	0.47 (0.04 to 0.89)

Note: Variables have been standardised to a mean of 100 and standard deviation of 10 to aid comparisons between measures. Sign for d adjusted so that positive treatment effects are positive and deterioration negative. Conventionally values of d are given as 0.20 = small, 0.5 = medium, 0.80 = large. Due to some missing data sample sizes are 1 n=43, 2 n=29, except 3 n=25 as many children did not attend school.

* trend towards significance p < 0.05 ** significant at alpha < 0.007 Bonferroni adjustment for seven tests

Table 5 shows:

- i) For the measure of CD/ODD there were clear and significant differences between pre and post assessment ($P < .001$). The effect size was in the medium range.
- ii) Child drug and alcohol use showed minimal improvement between the pre and post comparison; this difference was not statistically significant ($p > .05$) and this was reflected in an effect size close to nil.
- iii) Whilst parent rated child delinquent activities changed in the direction of improvement this was not statistically significant ($p > 0.05$) and the effect size was in the small range.
- iv) Child self-reported delinquent activities demonstrated a moderate reduction equivalent to a medium effect size and this displayed a trend towards statistical significance ($p < 0.05$).
- v) Teacher-rated CD/ODD behaviour improved to a modest degree equivalent to a small effect size and this was not statistically significant ($P > 0.05$)
- vi) With regards to ADHD related behaviours there was a clear reduction from the pre to the post assessment which was statistically significant ($p < 0.01$) with an effect size in the medium range.
- vii) A measure of parent-rated child depressive symptoms also demonstrated a statistically significant improvement from pre to post assessment ($p < 0.01$) which was equivalent to an effect size in the medium range.

Multivariate analysis of variance (MANOVA) was used to statistically combine data from the conduct problem measures and test for overall significance. For this analysis multiple imputation was used to estimate missing data points and allow the inclusion of data from all 58 participants. This analysis showed a degree of improvement across the six measures which was statistically significant and amounted to an effect size in the medium range (Wilk's $\lambda = 0.66$, $F(6, 52) = 4.36$, $p < 0.001$, multivariate eta squared 0.07^{14}). Given the presence of both missing and skewed data in the dataset the analysis was repeated with a) no imputed data and b) with truncated data to limit skew; both were statistically significant ($p < 0.01$) hence the first analysis (with imputed data) is reported here. These results show that despite minimal levels of change in child alcohol and drug use and small effects for parent reported child delinquent activities the combined outcome for conduct problem related measures indicated significant improvement from the beginning to the end of the treatment period.

¹⁴ For multivariate eta squared 0.01 = small, 0.06 medium 0.14 large

6.2 Outcomes pre to follow-up

Table 6 presents changes in conduct problem outcomes measured from pre-treatment to follow-up, twelve months after the commencement of treatment and approximately six months after the completion of treatment. ADHD and DEP data was not collected at follow-up.

Table 6 shows:

- i) For the measure of CD/ODD behaviours there was a clear difference between pre-treatment and follow up; this was statistically significant ($P < 0.01$) and in the order of a large effect size. As the effect size was larger than in Table 5 this shows that young persons continued to make gains over the post to follow-up period.
- ii) Child alcohol and drug use registered a slight improvement in the order of a small effect and this was not statistically significant ($p > 0.05$). The effect size was larger than in Table 5 indicating that minor gains were made over the post to follow-up period.

Table 6. Child anti-social behaviour and related outcomes pre-treatment to follow-up

Measure	Pre mean (SD)	Post mean (SD)	(df) t	p	d (95% CI)
CD/ODD ¹ (parent)	103.7 (9.4)	96.3 (9.3)	(45) 5.20	0.0001**	0.78 (0.36 to 1.2)
Child alcohol/drug use ¹	100.9 (11.8)	99.1 (7.8)	(45) 1.01	0.31	0.18 (-0.23 to 0.59)
Child delinquent activities (parent) ¹	102.0 (11.2)	98.0 (8.3)	(45) 2.14	0.04*	0.41 (-0.00 to 0.82)
Child delinquent activities (self report) ²	100.9 (6.9)	99.1 (12.4)	(31) 0.88	0.38	0.18 (-0.31 to 0.67)
CD/ODD (teacher) ³	103.4 (10.3)	96.6 (8.6)	(21) 2.88	0.009**	0.71 (0.10 to 1.31)

Sample sizes are 1 n=46, 2 n=32, 3 n=22 * trend towards significance $p < 0.05$ ** significant at alpha < 0.01 Bonferroni adjustment for five tests

- iii) Parent reported child delinquent activities showed a moderate degree of change in the region of a medium effect size and this showed a trend towards significance ($p < 0.05$). The effect size was larger than in Table 5 indicating that additional gains were made over the post to follow-up period.
- iv) Although child self-reported delinquent activities reduced over the pre to follow-up period this was not statistically significant ($p > 0.05$) and the effect size was in the small

range. This effect size was smaller than in Table 5 indicating that this measure showed a degree of deterioration in the post to follow-up period.

- v) Teacher-rated CD/ODD behaviours showed a clear reduction over the pre to follow-up period, this was statistically significant ($p < 0.01$) and amounted to an effect size in the medium to large range. This effect size was markedly larger than in Table 5 indicating that incremental gains in classroom behaviour were achieved over the post to follow-up period.

A multivariate analysis with imputed data was also performed. This analysis found a statistically significant difference between the pre and follow-up assessments in the region of a medium to large effect size (Wilk's $\lambda = 0.62$, $F(4, 54.) = 8.31$, $p < 0.0001$, multivariate eta squared 0.11). As with the previous MANOVA the analysis was repeated with a) no imputed data and b) truncated data to reduce skew; both analyses were statistically significant ($p < 0.02$) hence the first analysis was retained. This analysis shows that despite small effect sizes in two of four measures there was a significant overall reduction in conduct problems over the pre to follow-up period.

6.3 Placement and education outcomes

Table 7 indicates young person placement and education status at pre, post and follow-up assessment points.

Table 7. Placement and education changes before and after FFT

Measure	Pre	Post	Follow-up
Living with biological parent	74%	73%	74%
Living with caregiver	21%	24%	19%
Living with solo parent/caregiver	68%	62%	not assessed
Living independently	0%	0%	0%
Group home residential	0%	2%	4%
Secure CYF or YJ residence	0%	0%	2%
Prison	0%	0%	0%
Attending mainstream School	60%	67%	59%
Attending correspondence school	2%	7%	0%

Table 7 shows that:

- i) Living situations were relatively stable with approximately three quarters of young persons living at home at all assessment points.
- ii) In the great majority of cases where the young person was living with a biological parent this parent was in sole care.
- iii) Very few young persons transitioned to more restrictive placements such as group homes, CYFS secure residences or prison.
- iv) The proportion of young persons attending mainstream school remained relatively stable at a little over half.

Chapter 7

Family factors: Parenting and relationships

This chapter presents information relating to parenting and relationship factors which may underlie change in conduct problem outcomes.

7.1 Parenting and parent-child relationships

Table 8. Parenting practices and parent child relationship pre to post treatment

Measure	Pre mean (SD)	Post mean (SD)	(df) t	p	d (95% CI)
Parent-child relationship					
Parent-child bonding	99.0 (10.1)	101.0 (9.9)	(40) 1.46	0.15	0.19 (-0.24 to 0.63)
Parental involvement	100.7 (9.8)	99.3 (10.3)	(42) 0.77	0.45	-0.13 (-0.55 to 0.29)
Appropriate parenting					
Positive parenting	97.5 (11.0)	102.5 (8.3)	(42) -2.51	0.015*	0.52 (-0.09 to 0.94)
Non-violent discipline ¹	100.9 (9.6)	99.0 (10.4)	(41) 1.04	0.30	-0.19 (-0.62 to 0.24)
Inappropriate parenting					
Inconsistent discipline	102.0 (10.6)	97.7 (8.8)	(41) 2.82	0.007*	0.47 (0.04 to 0.91)
Poor supervision	99.6 (8.3)	100.4 (11.5)	(42) -0.44	0.66	-0.07 (-0.50 to 0.35)
Corporal punishment	98.2 (12.0)	101.8 (7.1)	(42) 1.76	0.09	0.36 (-0.07 to 0.78)
Lax discipline	99.6 (10.8)	100.4 (9.2)	(42) 0.46	0.65	0.08 (-0.34 to 0.51)
Over-reactive discipline	98.8 (10.5)	101.2 (9.4)	(42) 1.84	0.07	0.25 (-0.18 to 0.67)
Hostile discipline	101.4 (10.0)	98.6 (9.9)	(42) 1.61	0.11	0.29 (0.14 to 0.71)

Sign for d adjusted so that positive treatment effects are positive and deterioration negative.

Sample sizes are 43 except 1, 42 and 2, 41

* trend towards significance $p < 0.05$ ** significant at $\alpha < 0.005$ Bonferroni adjustment for ten tests

Table 8 summarises ten parenting related measures with tests of statistical significance and effect sizes calculated as Cohen's d. The table shows:

- i) There were clear differences in the use of inconsistent discipline and positive parenting pre to post with both displaying a trend towards statistical significance ($p < 0.05$) and effect sizes in the medium range.
- ii) Of inappropriate parenting practices corporal punishment, over-reactive discipline, lax discipline and hostile discipline all decreased with effect sizes in the small to medium range. Levels of poor supervision increased (deteriorated) marginally.
- iii) In terms of parent-child relationship measures; parent-child bonding increased with an effect size in the small range whilst parental involvement decreased with an effect size in the small range.
- iv) Outcomes for appropriate discipline practices were mixed with clear increases in positive parenting as noted above, whilst non-violent discipline decreased to a degree consistent with a small negative effect size.
- v) There was no obvious pattern in terms of which measures did or did not improve; inconsistent discipline was near-significant, but a similar measure, lax discipline, changed little. Positive parenting improved markedly but use of non-violent discipline decreased.
- vi) Overall, two of ten parenting measures showed near-significant improvements with effect sizes in the medium range, and eight of ten measures showed improvement to some degree. The overall average effect size was 0.18, a small effect.

As in the previous sections (see 6.2 and 6.3) multivariate analysis of variance was used to statistically combine data from the ten parenting measures and test for overall significance¹⁵. Multiple imputation was used to estimate missing data points and allow the inclusion of data from all 59 participants. This analysis showed a degree of improvement across the ten measures which was statistically significant and displayed an effect size in the small to medium range (Wilk's $\lambda = 0.71$, $F(10, 48) = 2.00$, $p = 0.05$, multivariate eta squared 0.03¹⁶). The analysis was repeated with truncated data to reduce skew and this was not significant ($p = 0.14$); similarly the analysis was also repeated with no imputed data and this was not significant ($p = 0.19$).

These results show that there are weak indications of a general improvement across the ten parenting measures but with only one of three analyses significant this conclusion is tentative.

¹⁵ Testing the effect of time in the analysis

¹⁶ For multivariate eta squared 0.01 = small, 0.06 medium 0.14 large

7.2 Parent conflict and relationships

Table 9 summarises nine measures of the caregivers' relationship with their child and partner. Sample size for the partner related questions was limited by the relatively few caregivers with a live in partner at pre-test and post-test, as a consequence the partner related measures have low statistical power.

Table 9. Parent conflict and relationships pre to post treatment

Measure	Pre mean (SD)	Post mean (SD)	(df) t	p	d (95% CI)
Primary Caregiver					
Verbal aggression to child ¹	101.5 (10.7)	98.5 (9.1)	(42) 1.77	0.08	0.30 (-0.12 to 0.73)
Physical aggression to child ¹	100.4 (12.1)	99.6 (7.4)	(42) 0.40	0.69	0.07 (-0.35 to 0.49)
Neglect ¹	99.7 (8.7)	100.3 (11.2)	(42) -0.27	0.78	-0.06 (-0.48 to 0.36)
Caregiver's partner					
Verbal aggression to child ³	101.7 (11.4)	98.3 (8.4)	(11) 1.33	0.21	0.34 (-0.47 to 1.14)
Physical aggression to child ²	103.4 (12.9)	96.6 (4.1)	(12) 2.52	0.04*	0.71 (-0.09 to 1.50)
Neglect ²	102.7 (13.9)	97.3 (0)	(12) 1.39	0.19	0.54 (-0.24 to 1.32)
Caregivers' relationship					
Violence to partner ³	102.6 (13.9)	97.4 (0)	(11) 1.30	0.22	0.53 (-0.29 to 1.34)
Violence from partner ²	99.9 (11.0)	100.1 (9.3)	(12) -0.09	0.93	-0.02 (-0.79 to 0.75)
Relationship quality ⁴	97.7 (11.4)	102.3 (8.5)	(7) 1.73	0.13	0.45 (-0.55 to 1.44)

Sample sizes are 1 n=43, 2 n=13, 3 n=12, 4 n=8. * trend towards significance p < 0.05

Table 9 shows:

- i) The primary caregivers' verbal aggression towards the child reduced somewhat from pre to post although this was not significant and the effect size was in the small to medium range.
- ii) The primary caregiver's physical aggression showed an improvement in the small range, and neglectful parenting increased slightly with an effect size in the small range. The great majority of caregivers reported nil aggression and neglect at the beginning of the study, hence the scope for improvement on these measures was limited.

- iii) The primary caregivers' partners achieved clear reductions in the use of physical aggression to the child; this showed a trend towards significance ($p < 0.05$) and an effect size in the large range. Whilst it was not statistically significant, neglectful parenting reduced markedly with an effect size in the medium range. Verbal aggression also reduced with an effect size in the small to medium range. More partners than caregivers were rated as using physical violence at the beginning of the study, hence there was greater scope for improvement on this measure.
- iv) The caregiver's violence to his/her partner reduced markedly from pre to post and, although this was not statistically significant, it was equivalent to a medium effect size. There was almost no difference in the level of violence received from the caregiver's partner from pre to post-test.
- vi) In a measure of relationship warmth and bonding there were clear differences pre to post which, although not significant, were equivalent to a medium effect size.
- vii) The partner related measures were completed by only eight to thirteen caregivers and were statistically underpowered.¹⁷ Focussing on effect size, which is not affected by sample size, it can be said that three of nine measures obtained medium effects or larger, four were in the small to medium range and two measures deteriorated slightly.

Due to the limited sample size for the partner and relationship variables a multivariate analysis was completed for the three primary caregiver variables only. Multiple imputation was used to estimate missing data points and allow the inclusion of data from all 60 participants. This analysis showed a degree of improvement across the three measures which was statistically significant and displayed an effect size in the small to medium range (Wilk's $\lambda = 0.86$, $F(3, 55) = 2.89$, $p < 0.04$, multivariate eta squared 0.05). The analysis was repeated with truncated data to reduce skew and this was also significant ($p < 0.05$), however the analysis was also repeated with no imputed data and this was not significant ($p = 0.36$). The difference was most likely due to the loss of statistical power when missing data was not imputed. This shows that there is moderate but partial evidence for a general improvement in primary caregiver relationships over the course of the study.

7.3 Delinquent peer association

Table 10 summarises two measures of the degree to which young persons associate with peers who engage in delinquent activities. Significance is tested using the t statistic and the effect size is Cohen's d.

¹⁷ Approximately 30 would be required for adequate power

Table 10. Delinquent peer association pre to post treatment

Measure	Pre mean (SD)	Post mean (SD)	(df) t	p	d (95% CI)
Association with delinquent peers (parent) ¹	100.3 (10.2)	99.7 (9.9)	(40) 0.36	0.72	0.05 (-0.38 to 0.49)
Friends' delinquent activities (child) ²	102.8 (10.7)	97.2 (8.6)	(24) 2.58	0.016*	0.57 (-0.0 to 1.14)

Sample size 1 = 41, 2 = 25. * significant $p < 0.025$, Bonferroni adjustment for two tests

Table 10 shows:

- i) According the parental report there was little change in the degree to which the child associated with delinquent peers. The difference in scores from pre to post was not significant and the effect size was very small.
- ii) According to the childrens' report there was a marked difference in the level of delinquent activities their friends engaged in over the period from pre to post test. This change was statistically significant ($p < 0.025$) and the effect size was in the medium range.

7.4 Maternal depression

Table 11 summarises the intensity of depressive symptoms amongst those reporting at least some depression. All respondents were female.

Table 11. Maternal depression severity

Measure	Pre mean (SD)	Post mean (SD)	(df) t	p	d (95% CI)
Maternal depression severity	101.8 (9.5)	98.2 (10.4)	(14) 0.80	0.43	0.36 (-0.36 to 1.08)

Sample size = 15

This table shows that the severity of depression as reported by those primary caregivers who reported some degree of depression at either pre or post-test decreased from pre to post to a degree equivalent to a small to medium effect size and this change was not statistically significant.

In order to ascertain if there was a change in the number of individuals reporting any depression from pre to post odds ratios were calculated and compared. Twenty of fifty nine caregivers reported at least some depression at the beginning of the study, whilst eight of forty two reported depression at post-test. The odds of reporting depression at the end of treatment were approximately half the odds of reporting depression at the commencement

of treatment; this change showed a weak trend towards significance ($p < 0.10$) (odds ratio 0.46, 95% CI 0.15-1.27, $\chi^2 = 2.70$, $p = 0.10$).

7.5 Exploratory multiple regression analyses

Multiple regression was used to test if change in one or more parenting or family variables significantly predicted change in parent reported CD/ODD and ADHD related behaviours. A preliminary regression analysis was conducted with respect to the background factors of age, sex, socio-economic status and ethnicity; these accounted for very little variability in CD/ODD and ADHD scores (7-10%) and were thus omitted from further analysis.

7.6 Conduct Disorder and Oppositional Defiant Disorder behaviours

Model selection criteria were used to identify the simplest model with the best predictive power¹⁸. The results of this regression indicated that three parenting/family factors predicted 41% of the variance and that this model was statistically significant ($F(4,38) = 9.08$, $p > 0.0001$, $r^2 = 0.41$, adjusted $r^2 = 0.36$). Coefficients for the parenting/family factors and tests of significance are presented in Table 12.

Table 12. Multiple regression predicting change in CD/ODD behaviours during FFT intervention

Predictor	Beta <i>B</i>	SE	Std β	t	p
parent-child bonding	0.42	0.14	0.39	2.94	0.006*
non-violent discipline	-0.26	0.10	-0.34	-2.67	0.01*
hostile discipline	0.14	0.10	0.18	1.46	0.15

Sample size 42 * trend $p < 0.05$ ** $p < 0.017$ Bonferroni correction for three tests. Signs are adjusted so that increases in positive measures and decreases in negative measures are both positive

Table 12 shows:

- i) Parent-child bonding (openness, trust and lack of conflict) had the strongest relationship with CD/ODD outcome (standardised beta 0.39) and this relationship was statistically significant ($p < 0.01$). Improvements in parent-child bonding were associated with reductions in CD/ODD behaviours.
- ii) Parental non-violent discipline had the second strongest relationship with outcome (standardised beta -0.34) and this was also statistically significant ($p < 0.01$). However the relationship was negative; greater reductions in CD/ODD behaviours were associated with decreases in the use of non-violent discipline.

¹⁸ AIC, Mallows Cp and adjusted r squared were used, see method section.

- iii) Although reduction in hostile discipline (shouting at, threatening the child) coincided with reduction in CD/ODD behaviours this was not statistically significant ($p>0.05$).
- iv) Non-violent discipline relates to parental use of appropriate consequences for misbehaviour such as time-out and removing privileges. The negative association between non-violent discipline and reduction in CD/ODD behaviours can be seen in light of an overall decline in non-violent discipline from pre-to post (see Table 1 above) and a negative correlation between improvement (reduction) in CD/ODD behaviours and improvement (increase) in non-violent discipline ($r = -0.45$, non-violent discipline decreases as CD/ODD behaviours decrease). One possible interpretation of this is that as CD/ODD behaviours reduced over time parents found a reduced need to use strategies such as time-out, hence use of these strategies also decreased over time.

7.7 Attention-deficit/Hyperactivity Disorder related behaviours

An identical analysis to that described in section 6.1 was performed to determine if one or more parenting or family variables significantly predicted improvement in parent reported ADHD behaviours. The results of this regression indicated that four parenting/family factors predicted 27% of the variance and that this model was statistically significant ($F(4,38) = 3.59$, $p>0.01$, $r^2 = 0.27$, adjusted $r^2 = 0.20$). Co-efficients for the parenting/family factors and tests of significance are presented in Table 12.

Table 13. Multiple regression predicting change in ADHD behaviours during FFT intervention

Predictor	Beta B	SE	Std β	t	p
hostile discipline	0.22	0.10	0.38	2.15	0.038*
verbal aggression	-0.24	0.12	-0.38	-2.01	0.051
parent child bonding	0.26	0.13	0.31	2.05	0.047*
non-violent discipline	-0.18	0.08	-0.30	-2.07	0.046*

Sample size 42 * trend $p < 0.05$ ** $p < 0.0125$ Bonferroni correction for four tests. Signs are adjusted so that increases in positive measures and decreases in negative measures are both positive

Table 13 shows:

- i) Hostile discipline had one of the strongest relationships with ADHD outcome (standardised beta 0.38) and this relationship showed a trend towards significance ($p < 0.05$). Reductions in ADHD behaviours were associated with reductions in hostile discipline practices.
- ii) Verbal aggression displayed a near-significant ($p > 0.05$) association with a negative coefficient (standardised beta = -0.38) indicating that after accounting for changes in parent-child bonding, non-violent discipline and hostile discipline, those parents reporting decreases in ADHD behaviours also reported increases in verbal aggression. As

the overall correlation between reduction in ADHD behaviours and reduction in verbal aggression was positive ($r=0.24$), beta has the opposite sign from what might be expected.

- iii) Increases in parent-child bonding were associated with reductions in ADHD behaviours and this relationship showed a trend towards significance ($p<0.05$).
- iv) Non-violent discipline showed a trend towards significance ($p<0.05$). Standardised beta was negative (-0.30) and the correlation between non-violent discipline and reduction in ADHD behaviours was also negative (-0.31) indicating that reductions in ADHD behaviour were associated with reductions in the use of non-violent discipline.

Considering the two regression analyses it is notable that they both independently identified improvements in parent-child bonding and reductions in both hostile discipline and non-violent discipline as components in models which predicted CD/ODD and ADHD outcomes, adding to the robustness of this finding.

The regression analyses in conjunction with sections demonstrates that those parenting measures which showed the strongest pre-post change overall (judging by effect size, see Tables 8 and 9) such as positive parenting, inconsistent discipline and corporal punishment were distinct from those measures showing the strongest correlation with improvement as identified in the regression analysis i.e. parent-child bonding, hostile discipline and non-violent discipline. This shows that the measures which responded the most to the FFT intervention (largest pre-post effect sizes) were distinct from those with the strongest relationship to improved behaviour (identified in the regression analysis).

Chapter 8

Outcome comparisons for Māori and Non-Māori and between groups

This chapter compares behaviour, parenting and relationship outcomes between Māori and Non-Māori, boys and girls, older and younger children and children with greater or lesser behaviour problems at the outset of treatment.

8.1 Comparison between Māori and Non-Māori: Anti-social behaviour and related outcomes

Table 14 presents effect sizes for anti-social behaviour calculated as Cohen's *d* for Māori and Non-Māori. Given the sample size statistical tests of potential differences between Māori and Non-Māori lack statistical power and have been omitted.

Table 14. Child anti-social behaviours and related outcomes. Effect size for Māori and Non-Māori, pre to post.

Measure	Māori <i>d</i>	95% CI	Non-Māori <i>d</i>	95% CI
CD/ODD ¹ (parent)	0.34	-0.29 to 0.98	0.75	0.16 to 1.33
Child alcohol/drug use ¹	0.30	-0.33 to 0.94	-0.09	-0.66 to 0.47
Child delinquent activities (parent) ¹	-0.03	-0.67 to 0.60	0.24	-0.32 to 0.81
Child delinquent activities (self report) ²	1.01	0.15 to 1.86	0.32	-0.35 to 0.98
CD/ODD (teacher) ³	-0.03	-0.87 to 0.80	0.42	-0.33 to 1.16
ADHD ¹	0.41	-0.23 to 1.05	0.54	-0.03 to 1.11
DEP ¹	0.17	-0.46 to 0.81	0.71	0.12 to 1.29

The sign for *d* has been adjusted so that positive treatment effects are positive and deterioration negative. Sample size: 1 *n*=19 Māori, 24 Non-Māori; 2 *n*=12 Māori, 17 Non-Māori; 3 *n*=11 Māori, 14 Non-Māori.

Table 14 shows:

- Non-Māori achieved larger effects than Māori in five of seven measures (CD/ODD, Child delinquent activities (parent report), CD/ODD (teacher), ADHD and DEP). Conversely Māori achieved larger effects than Non-Māori in two of seven measures (Child alcohol and drug use and Child delinquent activities (child report)).

- ii) Non-Māori achieved a slightly larger mean effect of 0.41 compared to 0.31 for Māori. Both values were small to medium effects. The difference between the two was equivalent to a Cohen's d of 0.10, which is half 0.20, the value conventionally described as a "small" effect.
- iii) The pairs of 95% confidence intervals all overlapped hence there was no clear indication (statistical tests aside) that any pair of d estimates were from different distributions and hence dissimilar. However, in the absence of statistical tests of sufficient power none of the pairs of d estimates are demonstrably equivalent¹⁹.

Table 15. Child anti-social behaviour. Effect size by ethnicity, pre to follow-up

Measure	Māori d	95% CI	Non-Māori d	95% CI
CD/ODD ¹ (parent)	0.68	0.05 to 1.29	0.90	0.32 to 1.48
Child alcohol/drug use ¹	0.08	-0.52 to 0.69	0.27	-0.29 to 0.82
Child delinquent activities (parent) ¹	0.14	-0.47 to 0.74	0.57	0.0 to 1.13
Child delinquent activities (self report) ²	0.99	0.16 to 1.80	-0.06	-0.70 to 0.57
CD/ODD (teacher) ³	0.45	-0.45 to 1.33	0.93	0.08 to 1.77

Sample size: 1 n=21 Māori, 25 Non-Māori; 2 n=13 Māori, 19 Non-Māori; 3 n=10 Māori, 12 Non-Māori. ADHD and DEP were not collected at follow-up.

Table 15 details effect sizes for antisocial behaviour outcomes as measured at follow-up. This table shows:

- i) Non-Māori achieved larger effects in four of five measures (CD/ODD, Child alcohol and drug use, Child delinquent activities (parent report) and CD/ODD (teacher)), whilst Māori achieved a larger effect for a single measure, Child reported delinquent activities.
- ii) Of these five measures the direction of the difference (whether Māori or Non-Māori did better) was the same as in Table 1 in four measures, suggesting a degree of stability in the pattern of outcomes from post-test to follow-up.
- iii) When averaged across the five measures the mean effects obtained were very similar; 0.47 and 0.52 respectively for Māori and Non-Māori, both medium effects. The difference between the two values was equivalent to a Cohen's d of 0.05, markedly less than the value of 0.20 conventionally termed a "small" effect.

¹⁹ In other words a test of the difference between the two estimates of d per measure would be non-significant, however this may be due to the low power of the test rather than any putative similarity of the two estimates.

- iv) As in the previous table the pairs of confidence intervals all overlap, which indicates a degree of similarity between the estimates. In the absence of statistical tests of sufficient power the pairs of estimates are not demonstrably equivalent.

8.2 Comparison between Māori and Non-Māori: Parenting practices

Table 16 presents effect sizes for ten parenting measures calculated as Cohen's *d* for Māori and Non-Māori separately.

Table 16. Parenting practices: Effect size by ethnicity, pre to post

Measure	Māori <i>d</i>	95% CI	Non-Māori <i>d</i>	95% CI
Parent-child relationship				
Parent-child bonding ²	0.14	-0.51 to 0.79	0.23	-0.35 to 0.81
Parental involvement	0.05	-0.59 to 0.68	0.19	-0.38 to 0.76
Appropriate parenting				
Positive parenting	0.75	0.09 to 1.40	0.29	-0.28 to 0.86
Non violent discipline ¹	-0.42	-1.07 to 0.24	0.02	-0.54 to 0.59
Inappropriate parenting				
Inconsistent discipline ¹	0.43	-0.23 to 1.09	0.51	-0.06 to 1.09
Poor supervision	0.10	-0.53 to 0.74	-0.18	-0.74 to 0.39
Corporal punishment	0.43	-0.21 to 1.07	0.30	-0.26 to 0.87
Lax discipline	0.0	-0.63 to 0.63	0.15	-0.41 to 0.72
Over-reactive discipline	0.28	-0.36 to 0.91	0.22	-0.35 to 0.78
Hostile discipline	0.24	-0.40 to 0.88	0.33	-0.24 to 0.90

The sign for *d* has been adjusted so that positive treatment effects are positive and deterioration negative. Sample size: 1 *n*=18 Māori, 24 Non-Māori; 2 *n*=18 Māori, 23 Non-Māori. Sample size for all others *n*= all 19 Māori, all 24 Non-Māori.

Table 16 shows:

- i) In terms of measures related to the parent-child relationship in both cases Non-Māori obtained the better outcome. All effects were in the small range with an average of 0.09 for Māori and 0.21 for Non-Māori. The magnitude of the differences in the estimates were themselves equivalent to small effects (*d* = 0.09, 0.14).
- ii) With regards to appropriate parenting Māori obtained the larger effect for positive parenting whilst Non-Māori obtained the larger effect for non-violent discipline. The differences between the effects obtained for each measure were in the medium range. Despite the variability in the effects obtained the mean effects across the two measures were very similar being 0.16 for Māori and 0.15 for Non-Māori, both small effects.

- iii) In terms of inappropriate parenting, outcomes were evenly split with Non-Māori obtaining larger effects for lax, inconsistent and hostile discipline whilst Māori obtained larger effects for poor supervision, corporal punishment and over-reactive discipline. Overall outcomes slightly favoured Māori who obtained an effect size of 0.25 averaged across the six measures compared to 0.22 for Non-Māori, both small effects. The difference between the two average values was an effect size of 0.03, a negligible value.
- iv) All pairs of confidence intervals overlapped hence there was no clear indication that any pair of d estimates were from different distributions and hence dissimilar. However as above, in the absence of statistical tests of sufficient power the pairs of estimates are not demonstrably equivalent.
- v) When averaged across all ten parenting measures, outcomes for Māori and Non-Māori are very similar with mean effect sizes of 0.20 and 0.21 respectively, both small effects.

8.3 Comparison between Māori and Non-Māori: Parental conflict and relationships

A number of measures of adverse family relationships were collected. Table 17 presents these measures with Cohen's d calculated for Māori and Non-Māori.

Table 17 shows:

- i) For the primary caregiver all three measures favoured Māori, although effect sizes for neglectful parenting were both negative. The average effects for Māori and Non-Māori respectively were 0.26 and -0.11, small or negative effects; the difference between the two was itself an effect of 0.37 which is a small to medium effect.
- ii) In terms of the primary caregiver's partner all three measures also favoured Māori. Some of the differences were considerable, for instance for both verbal aggression and physical assault the difference between the two effect sizes were themselves each equivalent to a large effect. The average effects for Māori and Non-Māori respectively were 1.10 and 0.26 and the difference was a large effect of 0.84. However these results must be interpreted with caution as they are based on sample sizes of less than ten per group²⁰.
- iii) With regards to the partner relationships Māori obtained the larger effect for one measure and Non-Māori for the two remaining measures. There were some very large differences, for instance relationship quality yielded a large negative effect (deterioration) for Māori and a medium to large positive effect for Non-Māori. The average effects for Māori and Non-Māori respectively were -0.16 and 0.49, a negative and a medium effect. The difference was a medium to large effect of 0.65. These results were also based on a very small sample and must be interpreted with caution

²⁰ Only a minority of caregivers reported having a live-in partner.

- iv) All pairs of confidence intervals overlap and as above there are no clear indications of difference, however the pairs of estimates are not demonstrably equivalent.
- v) When averaged across all ten measures, Māori achieved a somewhat larger effect than Non-Māori with means of 0.40 and 0.21, medium and small effects respectively. The difference between the two values was equivalent to a Cohen's d of 0.19, a small effect.

Table 17. Parent conflict and relationships: Effect size by ethnicity

Measure	Māori d	95% CI	Non-Māori d	(95% CI)
Primary caregiver				
Verbal aggression ¹	0.59	-0.06 to 1.23	0	-0.56 to 0.56
Physical aggression ¹	0.24	-0.40 to 0.87	-0.26	-0.83 to 0.30
Child neglect ¹	-0.05	-0.69 to 0.58	-0.07	-0.63 to 0.50
Caregiver's partner				
Verbal aggression ²	1.11	-0.26 to 2.43	-0.19	-1.23 to 0.87
Physical assault ³	1.55	-0.07 to 2.96	0.47	-0.53 to 1.46
Child neglect ³	0.63	-0.66 to 1.89	0.50	-0.50 to 1.49
Caregivers' relationship				
Violence to partner ⁴	0.82	-0.92 to 2.46	0.67	-0.29 to 1.61
Partner violence ⁵	-0.53	-1.93 to 0.90	0.10	-0.82 to 1.02
Relationship quality ⁶	-0.77	-0.95 to 2.41	0.69	-1.95 to 0.61

Sample size: 1 n= 19 Māori, 24 Non-Māori; 2 n=5 Māori, Non-Māori 7; 3 n=5 Māori, 8 Non-Māori; 4 n=4 Māori, 9 Non-Māori; 5 n=3 Māori, 9 Non-Māori; 6 n=3 Māori, 5 Non-Māori.

* trend towards significance, p <0.05 ** significant p<0.005 Bonferroni adjustment for nine tests

8.4 Comparison of outcomes between boys and girls

A similar analysis was performed with regards to outcomes for boys in comparison to girls. Girls obtained larger effects in five of seven outcomes at post-test with means of 0.60 (medium to large) compared to 0.26 (small to medium) for boys. At follow-up girls obtained larger effects in four of five measures with a mean effect size of 0.69 (medium to large) compared to 0.42 (small to medium). Although statistical tests could not be performed as noted above, girls appeared to be obtaining distinctly larger effects than boys.

8.5 Comparison of outcomes between older and younger

This analysis was repeated with regards to the age of young persons in the study. As the average age was thirteen and a half participants were divided into those thirteen and

younger (range 9-13 years) and those fourteen and older (range 14-17 years) at the beginning of the programme. Older participants obtained markedly larger effects in all seven outcomes at post-test with a mean of 0.66 (medium to large) compared to 0.10 (small) for older and younger respectively. At follow-up older participants maintained an advantage with a mean effect size of 0.54 (medium) compared to 0.26 (small to medium) for younger children. Thus older youth obtained noticeably greater medium to large effects overall in comparison with the small to medium effects obtained by younger children.

8.6 Comparison of outcomes for greater and lesser severity of behaviour at intake

The severity of problems at intake has been shown to be a predictor of outcome in a number of meta-analyses of interventions for child conduct problems. For instance Menting et al. (2013) in their analysis of 50 Incredible Years parent behaviour management training studies found that the strongest predictor of the treatment effect size for parent ratings was the pre-treatment intensity of child problem behaviour. In order to investigate this issue all measures were divided at the median pre-treatment score and outcomes for those young persons with pre scores larger than the median (for ease of reference the "severe" group) were compared with those with pre-test scores less than the median (the "mild" group)²¹.

For all measures at all time intervals (post-test, follow-up) the severe group obtained larger effect sizes than the mild group. For the measures CD/ODD and ADHD effect sizes were all positive with the mild group obtaining medium effect sizes at post and follow-up (range 0.46-0.66) whilst the severe group obtained large effect sizes (range 0.89-1.36). For parent and young person rated delinquent activities as well as parent rated child depression, at both post-test and follow-up the mild group obtained negative effect sizes (deterioration) in the small to medium range (-0.45 to -0.06) whilst the severe group obtained small to large positive effects (range 0.28-1.23). The teacher rating of CD/ODD behaviour displayed the same pattern with a negative effect size for the mild group at post-test but a medium effect for the severe group. Effects at follow-up were both positive but with a very large effect for the severe group ($d=1.71$). Effect sizes for the child alcohol and drug measure were negative for the mild group at post and follow-up but with small positive effects for the severe group. A notable consideration is that all those measures for which the mild group obtained negative effect sizes indicative of deterioration included a sizeable number of zero raw scores at pre-test which, as there are no negative raw scores, cannot improve. The negative effects obtained may therefore be in part due to a floor effect whereby if "perfect" scores are obtained initially, improvement is mathematically impossible.

²¹ An alternative strategy of comparing those with pre-test scores of plus or minus one standard deviation or more resulted in very small sample sizes.

Chapter 9

The influence of Māori identity

This chapter evaluates the role of a strong Māori identity in terms of pre-treatment levels of conduct problems and treatment progress.

9.1 Identity and outcome

There is no doubt that Māori are disadvantaged compared to Non-Māori in terms of a range of social, economic and health related outcomes. In this context the hypothesis that a strong Māori cultural identity serves as a protective factor with respect to undesirable social outcomes has been the subject of debate and analysis (Marie, Fergusson & Boden, 2008, 2009; Smith, 1999).

In order to ascertain if Māori cultural identity had an influence on pre-treatment anti-social behaviour and treatment outcomes Māori primary caregivers were first grouped into those identifying as solely Māori and those identifying as Māori and some other ethnicity. Māori participants were then asked about their cultural knowledge and participation via a series of items relating to knowledge of Te Reo and Māori protocol, attendance of cultural events such as unveilings and powhiri and the individual's satisfaction with his/her Māori identity. These items were aggregated to create a "cultural engagement score". This measure was derived from an instrument used in the Christchurch Health and Development Study.

Table 18. Child anti-social behaviour and related outcomes. Effect size for Sole-Māori and Māori-other ethnicities, pre to post.

Measure	Sole-Māori d	95% CI	Māori-other d	95% CI
CD/ODD ¹ (parent)	0.28	-0.57 to 1.12	0.27	-0.73 to 1.25
Child alcohol/drug use ¹	0.16	-0.68 to 0.99	0.53	-0.50 to 1.53
Child delinquent activities (parent) ¹	-0.40	-1.24 to 0.46	0.26	-0.74 to 1.24
Child delinquent activities (self report) ²	1.21	-0.18 to 2.51	0.66	-0.57 to 1.83
CD/ODD (teacher) ³	-0.15	-1.07 to 0.78	0.80	-1.03 to 2.49
ADHD ¹	0.24	-0.61 to 1.07	0.40	-0.61 to 1.39
DEP ¹	0.33	-0.52 to 1.17	-0.19	1.17 to 0.80

Sample size: 1 n=11 Sole-Māori 8 Māori-other, 2 n=6 Sole-Māori, 6 Māori-other, 3 n=9 Sole-Māori, 3 Māori-other.

Cultural engagement scores were significantly different for each group with Sole-Māori reporting a greater level of participation and cultural knowledge ($t(22) = 5.50$ $p < 0.001$). As such the hypothesis that cultural identity serves as a protective factor suggests that both pre-treatment levels of anti-social behaviour and outcomes post treatment should favour Sole-Māori when compared with those of Māori-other identity.

Effect sizes were calculated for each group for anti-social and related outcomes at post-test and follow-up.

Table 19. Child anti-social behaviour and related outcomes. Effect size for Sole-Māori and Māori-other ethnicities, pre to follow-up.

Measure	Sole-Māori d	95% CI	Māori-other d	95% CI
CD/ODD ¹ (parent)	0.48	-0.32 to 1.26	0.66	-0.56 to 1.83
Child alcohol/drug use ¹	0.49	-0.31 to 1.28	0.62	-0.60 to 1.78
Child delinquent activities (parent) ¹	-0.21	-0.98 to 0.57	0.91	-0.38 to 2.13
Child delinquent activities (self report) ²	1.0	-0.13 to 2.09	0.79	-0.60 to 2.09
CD/ODD (teacher) ³	0.98	-0.23 to 2.13	-0.47	-2.08 to 1.23

Sample size: 1 $n=13$ Sole-Māori 6 Māori-other, 2 $n=8$ Sole-Māori, 5 Māori-other, 3 $n=7$ Sole-Māori, 3 Māori-other.

Tables 18 and 19 show:

- i) Those of Māori-other identity achieved better outcomes than Sole-Māori in four of seven measures at post-test and three of five measures at follow-up. These proportions (four of seven, three of five) are not clearly different from what would be expected by chance.
- ii) Mean effect sizes were somewhat better for Māori-other at post-test ($d = 0.24, 0.39$) and slightly better for Sole-Māori at follow-up ($d = 0.55, 0.50$).
- iii) Overall this analysis was inconclusive; the hypothesis that Sole-Māori would achieve the better outcomes was not supported and if anything Māori-other yielded the better results. However these differences were not clearly shown in the data or demonstrated statistically.

9.2 Identity and initial symptom severity

Levels of anti-social and related outcomes were compared across ethnicities at the commencement of the study. Data for Non-Māori is included for comparison.

Table 20. Child anti-social behaviour and related outcomes. Standardised scores (mean 100 SD 10) at pre-test for Sole-Māori, Māori-other and Non-Māori ethnicities

Measure	Sole-Māori mean (SD)	Māori-other mean (SD)	Non-Māori Mean (SD)
CD/ODD ¹ (parent)	100.72 (12.0)	101.1 (8.40)	99.30 (9.52)
Child alcohol/drug use ¹	98.97 (9.78)	103.4 (13.0)	99.6 (9.31)
Child delinquent activities (parent) ¹	100.0 (16.0)	102.2 (8.04)	99.40 (5.73)
Child delinquent activities (self report) ²	102.26 (13.28)	98.10 (8.44)	99.68 (9.34)
CD/ODD (teacher) ³	96.50 (7.7)	99.90 (10.92)	101.80 (10.64)
ADHD ¹	101.95 (11.10)	105.91 (9.85)	97.31 (8.69)
DEP ¹	99.94 (11.87)	96.67 (10.19)	101.95 (8.96)

Sample size: 1 n=17 Sole-Māori 9 Māori-other 33 Non-Māori; 2 n=8 Sole-Māori 6 Māori-other 21 Non-Māori; 3 n=12 Sole-Māori 7 Māori-other 25 Non-Māori.

Table 20 shows:

- i) Of the seven measures Non-Māori obtained the “best” (lowest) mean score for three measures, Māori-other two and Sole-Māori two, hence scores overlapped and no ethnicity exhibited consistently lower levels of problematic behaviours at the outset of the study.
- ii) Of the seven measures Māori-other obtained the “worst” (highest) mean score for four measures, Sole-Māori one and Non-Māori two, hence there was a modest tendency for Māori-other participants to exhibit higher levels of problematic behaviours at the outset of the study.
- iii) Multiple t-tests were used to compare scores on each measure for Māori-other vs. Sole-Māori; none were significant ($p>0.33$).
- iv) When comparing Sole-Māori with Māori-other the former obtained the lowest mean score for five of seven measures, hence there was a modest, but not statistically verified, tendency for Sole-Māori to exhibit a lesser degree of problematic behaviour and other issues prior to the commencement of the FFT study.
- v) These outcomes are consistent with the hypothesis that Sole-Māori identity confers a reduced risk of conduct problems prior to intervention when compared to those of Māori-other ethnicity.

Overall the analysis of treatment effects was quite inconclusive with neither ethnicity showing a clear advantage. The analysis of anti-social behaviour at the commencement of the study was consistent with the hypothesis of an identity related protective factor, however this did not amount to strong or statistically verified evidence for this hypothesis.

Chapter 10

Parent satisfaction

This chapter presents information relating to the degree to which parents were satisfied with various aspects of the FFT intervention. Māori parents were asked about their satisfaction with FFT from a Māori cultural perspective.

10.1 Satisfaction with aspects of FFT

Parental satisfaction with FFT was generally high. Table 21 shows that therapist characteristics were highly rated with 80% or more of parents reporting being “very much” satisfied with the therapist’s timeliness, appointment keeping and support. Almost 80% of parents were “very much” satisfied with the competence of their therapist.

Table 21. Parent satisfaction with aspects of FFT: Percentage of parents’ responses

How satisfied were you with...	Not at all	A little	Moderately	Very much
The way FFT was explained to you before you started FFT	5%	12%	21%	63%
The way FFT was introduced to you by your therapist during the first few FFT sessions	7%	2%	28%	63%
The convenience of the session times	5%	5%	12%	73%
The timeliness of the therapist (turning up on time)	2%	2%	9%	86%
The therapist’s appointment keeping (not cancelling or missing sessions)	2%	2%	7%	88%
The overall value of FFT for your family	12%	5%	9%	74%
The impact of FFT on your son/daughter’s behaviour	16%	14%	14%	56%
The impact of FFT on your son/daughter’s ability to get on with the rest of the family	16%	14%	23%	46%
The competence of your therapist	5%	2%	14%	79%
The level of support given by your therapist	7%	0%	12%	81%
The extent to which you got the type of therapy that you wanted	9%	9%	14%	67%

Sample size 43 Some rows may not sum to 100 due to rounding error

Approximately three quarters of parents were “very much” satisfied with the overall value of FFT for their family and two-thirds felt that FFT was the right sort of therapy for their family. In contrast, only approximately half were “very much” satisfied with FFT’s impact upon their son or daughter’s behaviour and ability to get on with the rest of the family. This suggests that whilst parents felt that FFT was a valuable intervention and gains were made behavioural and relationship difficulties were not always entirely resolved at the completion of the intervention. The presentation of FFT to families appeared adequate with two-thirds or more of parents being “very much” satisfied with the way FFT was introduced or explained to them.

10.2 Satisfaction for Māori

Māori parents were very satisfied with the cultural knowledge and respect shown by their therapist with 80-90% reporting that they were “very much” satisfied (see Table 22). A similar proportion of parents felt that FFT was an appropriate intervention for Māori whānau.

Table 22. Parents of Māori descent: Satisfaction with the cultural sensitivity of FFT

How satisfied were you with...	Not at all	A little	Moderately	Very much
The respect shown by your therapist for your culture or tikanga	10%	0%	0%	90%
The level of knowledge shown by your therapist of your culture or tikanga	10%	0%	10%	80%
The suitability of FFT as a therapy for Māori Whanau	10%	0%	10%	80%

Sample size n=20

Chapter 11

Treatment fidelity and outcomes

This chapter investigates the relationship between therapist treatment fidelity and outcomes.

The fidelity with which an intervention is delivered is a key issue in research. In particular poor treatment fidelity may impair outcomes for an otherwise effective intervention (Henggeler et al., 1997). A valid evaluation of an intervention can only be made if the intervention is delivered correctly.

11.1 Measures and procedure

The FFT model includes systematic monitoring of implementation and therapist competence. Each therapist is assessed quarterly by the programme supervisor in terms of 36 skills relating to five phases or aspects of FFT treatment; engagement, behaviour change, generalisation, general FFT skills and relationship skills. Each therapist is also rated weekly in terms of *dissemination fidelity* (tangible tasks such as attendance at meetings and completion of paperwork) and *fidelity* (overall fidelity to the FFT model) with respect to a specific case. The quarterly assessment rates skills from 0 to 7 and averaged within the five domains to yield five scores, and the weekly assessment rates skills from 0 to 6.

In order to correlate fidelity measures with client outcomes for the present study two methods were used. In the first analysis the therapist's quarterly scores were assigned to every participant treated by that therapist over that period. Where a participant's treatment spanned two quarterly reports the scores were averaged and then assigned to the participant. In this way each participant was assigned five scores representing the fidelity of the treatment they received.

Multiple correlations were calculated between the five fidelity scores per participant and the seven post-treatment and five follow-up conduct problem related outcomes. An exploratory approach was taken and alpha was set at a nominal $p=0.05$ per contrast.

The second analysis was similar except that the fidelity score assigned to each participant was the average of the weekly fidelity ratings assigned to that participant. Some participants were not rated by the programme supervisor.

11.2 Results

In the first analysis correlations between the fidelity measures and outcomes were generally small and none were significant ($P>0.05$). The mean correlation between the five fidelity measures and the seven post-treatment outcomes was - 0.03 (range -0.36 to 0.27). The mean correlation between the five fidelity measures and the five follow-up measures was -

0.04 (range -0.21 to 0.27). These near-nil correlations suggest almost no relationship between the fidelity assessments and outcomes, however as several participants may be assigned the same fidelity scores (see above) there may have been limited variability in fidelity scores and hence low statistical power.

In the second analysis two fidelity ratings reflecting the quality of the therapist's management of specific cases were correlated with the conduct problem outcomes for those cases. Correlations were again generally small; the mean correlation between the *fidelity* measure and the seven post-treatment outcome measures was -0.08 (range -0.32 to 0.26) and for the five follow-up measures -0.12 (range -0.37 to 0.07). Two of these 12 correlations were significant ($p < 0.05$), however the correlations were negative; higher fidelity scores obtained poorer results. The mean correlation between the *dissemination fidelity* measure and the seven post-treatment outcome measures was -0.02 (range -0.16 to 0.15) and for the five follow-up measures -0.08 (range -0.21 to 0.09). No correlations were significant.

The present findings have four possible interpretations:

Firstly, for whatever reason, the treatment fidelity measures used may not adequately capture variations in treatment fidelity. The measures may be insensitive or inaccurate.

Secondly, the measures are accurate but statistical power is limited. Sample sizes were small and hence the correlations were underpowered. This is certainly correct; with 44 observations in the second analysis, power to detect a small effect would be 0.16, very low.²² However, power should not affect the sign of the correlations and 10 of 14 dissemination fidelity and fidelity correlations at post-treatment were negative and 8 of 10 at follow-up. For the quarterly measures 21 of 35 correlations at post-test were negative and 17 of 25 at follow-up. Negative correlations indicate that better fidelity scores are associated with poorer outcomes.

The third possibility is that as fidelity ratings were generally high, the limited variability would limit the size of correlations. There is some evidence for this; for instance the mean score for dissemination fidelity was 5.4 and for general fidelity 4.3 in a 6 point scale.

A fourth possibility is that variations in outcome were related to other factors such as age, sex, ethnicity and initial severity of problem behaviours and that in the absence of an analysis controlling for these factors the influence of treatment fidelity cannot be detected.

²² Meaning that there is a 16% chance of rejecting the null-hypothesis where this hypothesis is true; adequate power is usually considered to be 0.80 or more.

Chapter 12

Summary and discussion

This pilot study was undertaken by Youth Horizons and the University of Otago to ascertain whether Functional Family Therapy (FFT) in New Zealand was achieving similar results as those achieved internationally. The study was conducted from January 2011 - November 2013 with the first FFT team established in New Zealand, at Youth Horizons Trust, Auckland.

The research design of this pilot was a single-group outcome study with pre-test, post-test (six months following commencement of FFT) and follow-up (12 months after commencement of FFT). Each assessment point involved a parent interview, a child questionnaire and a teacher questionnaire where the child was attending school. Fifty-nine young people with conduct problems and their families were involved in the study.

12.1 Key findings

In general, where similar measures are used, the present study obtained results comparable to those achieved in previous overseas FFT and family-based interventions.

At post-test small to medium effects were achieved over the seven conduct problem and related outcomes with effects ranging from 0.58 for parent reported CD/ODD behaviours to 0.02 for child alcohol and drug use. The average effect across seven measures was 0.34, a small to medium effect. Changes in three measures were statistically significant and one measure displayed a trend towards significance ($p < 0.05$). A multivariate analysis which aggregated the seven outcomes was also significant ($p < 0.003$), showing that notwithstanding some small effects the overall outcome was a significant reduction in conduct and related problems.

At follow-up, one year after the commencement of treatment, effect sizes were in the small to large range with effects ranging from 0.78, a large effect, for CD/ODD behaviours to 0.18, a small effect, for Child Alcohol and Drug use. The average effect across five measures was 0.47, a medium effect. At follow-up two conduct problem measures became statistically significant and one displayed a trend towards significance ($p < 0.05$). A multivariate analysis which aggregated the five measures was also statistically significant.

12.2 Summary of findings related to the research hypotheses

The findings of this FFT pilot study provide evidence in response to the following research questions:

Effectiveness. Does FFT, as delivered by a community agency in New Zealand, lead to beneficial changes from pre to post treatment in key conduct problem related outcomes immediately after the cessation of treatment?

Whilst outcomes were variable significant or near significant medium effects were obtained in five of seven measures including parent-rated Conduct Disorder/Oppositional Defiant Disorder behaviours.

Maintenance of gains. Where beneficial changes are achieved immediately after the cessation of treatment, to what extent are these gains maintained six-months after the cessation of treatment?

Gains were largely maintained with four of the five measures improving over the follow-up period and two measures achieving significant effects in the large range.

Equity of outcomes for Māori. Where beneficial changes are achieved by Non-Māori youth at post-treatment or follow-up, to what extent are these gains shared by Māori?

Outcomes for Māori and Non-Māori were broadly similar. There were no statistically significant differences between Māori and Non-Māori on any measure. Non-Māori outcomes were slightly better for conduct problems and almost identical for parenting. Non-Māori obtained slightly better outcomes for parent-child relationship and moderately better outcomes for caregiver relationship measures. Māori obtained moderately better outcomes for caregiver-child conflict measures.

The influence of culture. For Māori whānau, to what extent is their degree of immersion in Māori culture related to the initial level of conduct problem outcomes at the commencement of treatment and subsequent progress?

Parents who identified as being of Sole-Māori descent had a significantly greater involvement in aspects of Māori culture than those identifying as Māori-other descent. There were no consistent differences in outcomes between the two groups however there were indications that Sole-Māori had a lower level of conduct problems at the beginning of the study.

Characteristics of the population treated. What are the social, economic, cultural and demographic characteristics of those families receiving the FFT intervention?

Participants reported experiencing a pervasive and disparate range of social and economic difficulties including low household income and education, high rates of benefit dependence and single-parenthood, high levels of school non-attendance, police contact, substance use, depression and multiple adverse life-events such as illness, bereavement, family arguments and financial problems. Approximately half of the participating families were Māori and two-thirds of the participating young persons were boys.

Which, if any, changes in parent management of youth behaviour from pre to post treatment predict conduct problem outcomes?

Improvements in parent-child bonding and reductions in hostile discipline demonstrated the strongest associations with improvements in conduct disorder/oppositional defiant disorder and ADHD related behaviours. However the parenting measures which showed the greatest response to the intervention were positive parenting and inconsistent discipline. Therefore the parenting measures which showed the greatest improvement were not those most strongly associated with change in conduct problems.

Treatment fidelity. To what extent, if any, does therapist to therapist variability in treatment fidelity predict conduct problem outcomes?

This study did not detect any significant correlation between fidelity measures and outcome.

Client satisfaction – overall. To what extent are parents satisfied with the FFT intervention?

Client satisfaction was very high overall with an average of 70% of parents reporting that they were “very much” satisfied with the effectiveness and presentation of FFT, compared to an average of 8% who reported being “not at all” satisfied.

Client satisfaction – Māori. To what extent do Māori parents find the FFT intervention culturally acceptable?

Client satisfaction amongst Māori for issues relating to knowledge of and respect for Māori tikanga was high, with 80-90% reporting being “very much” satisfied and 10% reporting being “not at all” satisfied.

12.3 Comparison of findings with comparable intervention studies

The findings of the present study are compared with those achieved in previous international FFT studies, and then also with MST studies internationally and within New Zealand.

When interpreting effect sizes a complicating factor is that effects when measured pre-post (PvP), as in the present study, may sometimes exceed effect sizes measured as the difference between a treatment and a control group (TvC) at the end of the study. For instance if the control group is receiving an alternative treatment it may improve somewhat, diminishing the difference between the groups. In one survey pre-post effect sizes exceeded treatment-control effect sizes by 0.13 on average (Wilson & Lipsey, 2001). Whilst pre-post and treatment-control effect sizes can be compared, the slight upward bias in pre-post effect sizes must be taken into consideration.

Only two published studies of FFT cite Cohen’s *d* for behaviour; Graham et al. (2013) and White et al. (2013) obtained effect sizes of 0.64 and 0.41 respectively for conduct problems which when combined using meta-analysis techniques yield an overall medium effect of $d = 0.53$. As these effect sizes were measured pre-post they are directly comparable with the present study where the effect size of 0.55 was obtained for CD/ODD behaviour at post-test.

The outcome obtained in the present study was thus very similar to that obtained in previous published studies.

Three previous studies examined the influence of FFT on internalising symptoms such as anxiety and depression; Graham et al. (2013); Rohde, Waldron, Turner, Brody & Jorgenson (2014) and White et al. (2013). Effect sizes ranged from the very small to medium. When aggregated using meta-analysis techniques the overall effect size was $d = 0.32$, a small to medium effect. The outcome obtained in the present study, a medium effect of $d = 0.45$, was thus somewhat larger than the mean effect obtained in previous studies.

Four previous studies involved substance or alcohol using adolescents; Friedman (1989), Rohde et al. (2013), Slesnick & Prestopnik (2009) and Waldron, Slesnick, Brody, Turner & Peterson (2001). Outcome effect sizes ranged from 0.06 to 0.86; when aggregated using meta-analysis techniques the overall effect size was a medium effect of $d = 0.45$. As such the outcome achieved in the present study, the near-nil effect of $d = 0.02$, is clearly below that obtained in previous FFT studies.

A factor to consider here is that in all four of the studies above the participants either had a DSM-IV diagnosis of a substance use disorder, or their primary presenting issue was substance abuse. In contrast, all participants in the present study were screened using the DSM-IV criteria for substance abuse and only two participants met the criteria and fifty percent reported nil drug or alcohol use at pre-test. Thus the obtained results may reflect the low level of substance use amongst participants at the commencement of the study.

The remainder of the FFT studies reviewed in Chapter 2 cite odds ratios for official offending data as opposed to Cohen's d for CD/ODD behaviours and delinquency as used in the present study.²³ When odds ratios for offending are converted into Cohen's d equivalents the mean effect size for the seven FFT studies in Table 1 is 0.85, a large effect (range 0-1.52)²⁴. This clearly exceeds the small to medium effects obtained (on average) in the present study, however the applicability of the comparison is tenuous as it involves different statistics (Cohen's d vs. odds ratios) and different outcomes (official offending vs. parent, child & teacher reported behaviour). Where the present study has been contrasted with previous studies using common outcome measures the results have been largely similar.

FFT studies cite generally favourable follow-up recidivism rates up to five years post-treatment but again the data is in the form of odds ratios. In the present study the majority of measures improved over the follow-up period, however in the absence of behavioural data from almost all previous FFT studies and official offending data for the present study an exact comparison of follow-up outcomes cannot be made.

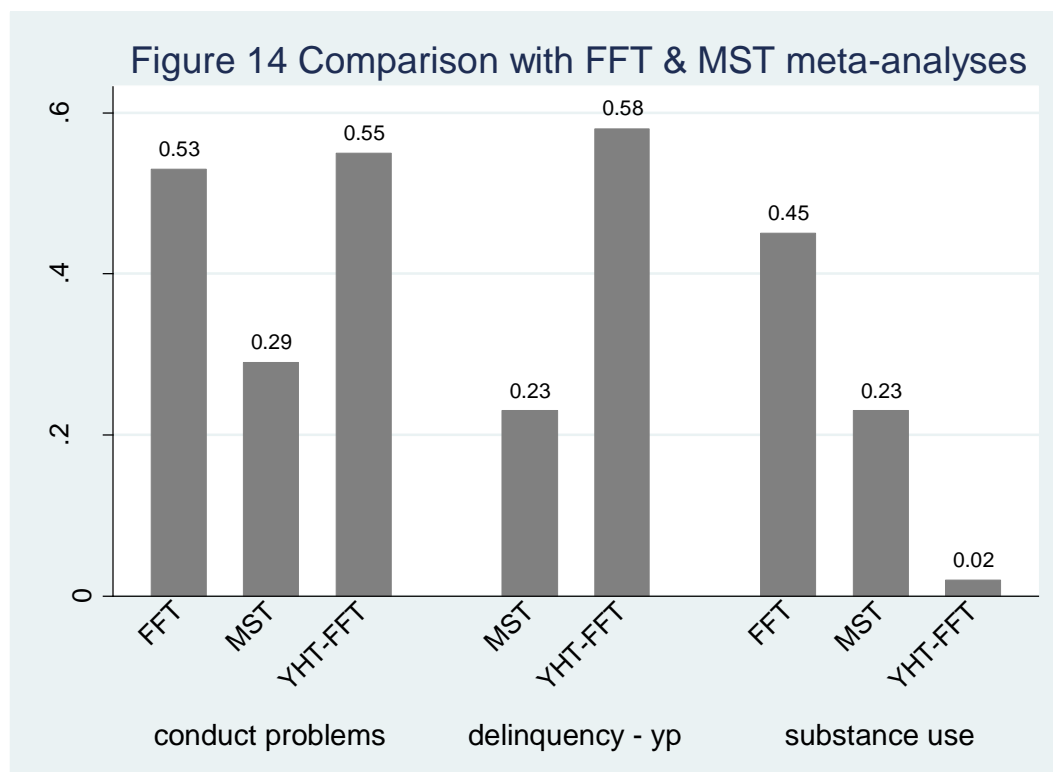
Multi-systemic Therapy (MST) is an appropriate comparison for examining the present study's findings as both FFT and MST are evidence based family interventions, both work in the main with adolescents rather than young children and there are some New Zealand studies. The MST studies, unlike most FFT studies, commonly cite outcomes which are

²³ Official offending data was unavailable for the present study

²⁴ OR to d formula cited in Farrington & Welsh (2003)

comparable with the present study such as self-report delinquency and parent-completed behaviour rating scales similar to the CD/ODD measure used in this study. Further context to the present findings is shown in Figure 14 which summarises outcomes in three key domains for the present study ("YHT-FFT"), the brief meta-analysis of FFT from chapter 2 ("FFT") and the most recent meta-analyses of MST; Stouwe, Asscher, Stams, Dekovic & Laan 2014 ("MST"). All variables except "YHT-FFT" represent the aggregated outcomes of multiple studies.

It can be seen that the effect size for parent-rated problem behaviour for the present study exceeds the mean effect sizes for the MST meta-analysis and is similar to that achieved by FFT. Similarly the effect size for young-person self-reported delinquency in the present study (delinquency - yp) was clearly well above the mean effect achieved across multiple published MST studies (no previous FFT studies have measured self-report delinquency). However, as previously noted, the effect size achieved in the present study for substance use was very small and clearly inferior to mean effects achieved by MST and previously by FFT.



As noted in Chapter 2 the two previous FFT studies conducted outside the United States, where the programme originated did not appear to suffer a disadvantage. This is in contrast to finding for MST where non-USA based studies have tended to obtain poorer outcomes (cf. Stouwe et al., 2014). Only one non-USA based study, Graham et al. (2013) based in Ireland, collected comparable data to the present study. Although a sample of one study provides little grounds for appraising the accuracy of the comparison, Graham et al. obtained slightly better results than the present study (e.g. d for conduct problems 0.64 vs. 0.55) and somewhat better results than comparable USA-based studies (e.g. 0.46 vs 0.46 & 0.05 see

Table 1). As such there was no clear evidence of a non-USA disadvantage for any study including this study.

The final point of general comparison is with New Zealand based studies. While the present study is the first study of FFT in New Zealand, there have been three previous replications of MST in New Zealand to our knowledge and these are of interest insofar as both FFT and MST are family based therapies which have been developed overseas (see Appendix 1.). All studies were single-group outcomes studies with no control group. Grace, McLean & Warren (2006) represented an evaluation of MST as used in the Reducing Youth Offending Programme. Referrals to this programme were via Youth Justice rather than CYF. The primary offending measure used a complex statistical modelling technique not directly comparable to any outcome cited here. Where pre and post measures were obtained conduct related effects ranged from -0.08 to 0.53. By and large, where the measures were similar, the present study recorded somewhat better outcomes e.g. self-reported delinquency 0.29 vs. 0.58. Russell (2008) is an unpublished PhD study where the author documented outcomes and treatment fidelity for a community-based MST team. Rating scale measures of offending changed little, however a measure of general problem behaviour obtained a medium to large effect size of 0.67, exceeding most outcomes in the present study. Curtis, Ronan, Heiblum & Crellin (2009) also evaluated community-based MST teams. This was a well conducted study however outcomes were modest with small effects ranging from $d = 0.16$ for offense severity to $d = 0.29$ for a measure of problem behaviours.

To conclude, although the comparison with a different intervention is at best partial, it can be seen that the outcomes obtained in the present FFT pilot study are at least on a par with those previously obtained in New Zealand replications of a similar family-based intervention.

12.4 Discussion of additional outcomes

Reduction in child attention problems. Reduction in child attention problems has not been previously reported in FFT or MST studies. The present study achieved a medium effect size of 0.45 for ADHD symptoms at post-test. This result falls within the range of effect sizes previously reported for non-pharmacological treatments for ADHD. For instance in a study of the Incredible Years parent training intervention Webster-Stratton, Reid & Beauchaine (2013) obtained effects ranging from 0.41 inattention to 0.98 for hyperactivity and in a New Zealand Incredible Years replication Sturrock & Gray (2013) reported an effect size of 0.55 for ADHD symptoms (pre to post). In a comprehensive recent meta-analysis of behavioural interventions for ADHD Fabiano et al. (2009) reported medium to large effects overall for behavioural interventions in the region of 0.67-0.83 for between group and pre-post designs. An important consideration is that FFT is not intended as a specific intervention for ADHD and any improvement in ADHD related behaviours is secondary to a primary focus on delinquency. As behavioural intervention is a common and generally effective treatment for ADHD we can surmise that behavioural strategies used within the "behaviour change" phase of FFT have impacted upon ADHD related behaviours.

Teacher reported conduct problems. The current study achieved variable effects for teacher reported conduct problems with a small effect of 0.21 at post-test and a medium to large effect of 0.71 at follow-up. Although there are no analogous FFT or MST studies with which to compare, outcomes for parent training interventions with no teacher training component generally fall within the range achieved here. Effect sizes at school for interventions with no specific classroom component are generally more modest than effects achieved in the home as rated by parents. For instance in one study Incredible Years parent training obtained an effect size of 0.22 for child conduct problems at school (Webster-Stratton, Reid & Hammond, 2004). Sturrock & Gray (2013) in a New Zealand study cited teacher reported effects of 0.17 at post-test and 0.30 at follow-up (pre-post). In a meta-analysis of all available Incredible Years studies Menting, Orobio de Castro & Matthys (2013) report a mean effect size of 0.13 for teacher reported conduct problems.

Outcome by gender. Girls obtained markedly larger effects than boys at both post-test and follow-up with the difference between boys and girls being equivalent to a small to medium effect. Results from the intervention literature are variable; Baldwin Christian, Berkeljon, Shadish & Bean (2012) and Fossum, Handegard, Martinussen & Morch (2008) in meta-analyses of family therapies and interventions for disruptive behaviour respectively both found that girls obtained poorer outcomes than boys. Ogden (2006) in a study of MST found that girls improved more in terms of externalising behaviour and delinquency whilst boys improved more in terms of internalising behaviours.

Outcome by age. Older children obtained larger effect sizes on average than younger. This is in contrast to the findings of some reviews; Fossum et al. (2008) found that younger children obtained better outcomes than older across the range of psychosocial interventions for disruptive behaviour however the more apt comparison may be with Cognitive behavioural interventions where increasing age is associated with better outcomes (McCart, Priester, Davies & Azen, 2006). Insofar as the process of FFT involves dialogue and rational discussion it resembles Cognitive behaviour therapy and like CBT may be more suited to older and more cognitively mature children and youth. Conversely behavioural intervention, where younger children tend to obtain better outcomes (McCart et al. *ibid.*), is a secondary component in FFT.

Problem severity. Children with a greater intensity of problem behaviour at the commencement of FFT obtained larger effects on average compared with children who exhibited a milder degree of problem behaviour. This is a not uncommon finding in the disruptive behaviour literature, for instance Menting et al. (2013) in their analysis of 50 Incredible Years parent behaviour management training studies found that the strongest predictor of the treatment effect size for parent ratings was the pre-treatment intensity of child problem behaviour. To some extent this trend is purely mathematical; higher initial scores have a greater distance to fall to an average or below average score and this will translate into larger effect sizes. It may also be the case that more severe behaviours engender a greater motivation on the part of parents and professionals to change the behaviour.

Placement outcomes. Despite many young persons being at risk of out-of-home placement only six percent transitioned to more restrictive settings such as group homes, secure

residences or prison. Approximately three quarters of young persons were living with their biological parents at the commencement of the study and this proportion remained fairly constant across all three assessment points. As family preservation and the retention of children and young persons within the family is a key value for social services in New Zealand this is a significant finding which shows that FFT by and large is achieving this goal.

Educational outcomes. Over a third of young persons were not attending mainstream school at the onset of the study and this proportion decreased slightly over the course of FFT and increased slightly over the follow-up period. Any general tendency for young persons to return to school is likely to have been counter-balanced by older youth reaching school-leaving age. As such FFT did not significantly alter school attendance rates.

The second largest group at follow-up after those attending school or education was those young persons who had left school but were not studying or employed (15%); whilst regrettable this figure must be interpreted in light of the youth unemployment rate, which has been 24-25% of 15-19 year olds in recent years.²⁵

Outcomes for Māori. Although Non-Māori obtained larger effects than Māori for the majority of child and parenting outcome measures the differences between the two were equivalent to a small effect or less on average and sample sizes were too small to test for statistical significance. In the domain of parent conflict effects sizes favoured Māori and the difference between the two was more pronounced, a small to medium effect. Effect sizes relating to the caregivers' partner's behaviour favoured Māori by a large margin and effects for the caregiver's relationship favoured Non-Māori by a similar margin however the sample sizes for these two categories were very small and the effects yielded must be interpreted with caution. This finding is consistent with other New Zealand studies where comparative data has been collected; for instance Sturrock & Gray (2013) in their study of the Incredible Years programme found larger effects for non-Māori which were significant for child behaviour at follow-up but not at post-test and not for parenting practices and relationships.

For Māori, it has been hypothesised that a strong identity as Māori is an important mitigating factor with respect to a range of adverse social outcomes including crime, substance use and mental health. In the present study those parents reporting Sole-Māori ethnicity reported higher levels of engagement in and knowledge of Māori cultural practices than those of Māori-other ethnicity. Previous research has shown that Māori-other have elevated levels of mental health difficulties and offending (compared to Non-Māori) which cannot be explained by socio-economic or family adversity factors; this finding is consistent with the hypothesis that Māori-other suffer from an "afflicted cultural identity" (Marie, Fergusson & Boden 2008, 2009). On this basis it might be hypothesised that Sole-Māori would experience better outcomes than Māori-other; this was not found, conduct problem outcomes marginally favoured Māori-other at post-test and slightly favoured Sole-Māori at follow-up. However, there was a modest but not statistically significant trend for participants of Sole-Māori descent to display a lower level of conduct problems at the commencement of treatment and this is consistent with previous research. For instance, Marie, Fergusson & Boden (2009) found Māori-other to have higher levels of official and self-reported offending

²⁵ Statistics New Zealand, NZ social indicators.

than Sole-Māori prior to any statistical adjustment for socio-economic and family adversity factors.

12.5 Moderating factors

Parent discipline practices. There is considerable evidence that parental discipline practices are linked to child conduct problems (Frick et al., 1992; Shelton, Frick & Wootton, 1996). Those aspects of parenting most commonly identified in this context include inconsistent discipline, failure to use positive discipline strategies and excessive use of corporal punishment. A number of parent behaviour management training studies have shown that parenting practices are amenable to change and that where the study shows improvements in child behaviour in the medium to large range, effect sizes for parenting practices tend to be in the medium range. (e.g. Beauchaine et al., 2005; Bjorknes, Kjobli, Manger & Jakobsen 2012; Sanders, Baker & Turner, 2012; Schaffer, Lindheim & Kolko 2013; Sturrock & Gray, 2013). In the present study significant improvements were achieved for positive parenting and inconsistent discipline however effect sizes for the remaining eight measures were generally small. One possible reason for the lower effect sizes achieved in this study is that parent training interventions specifically target parenting practices whereas FFT focusses on a broader rubric of relational and motivational patterns within the family. A second reason is that research on parenting practices has largely been in the context of parent behaviour management training interventions for children under 10, as opposed to the adolescent age-group treated by FFT; there is evidence that interventions for adolescents are less effective than interventions for younger children hence the present outcomes for parenting may reflect the relative performance of interventions for adolescents (Dishion & Patterson, 1992). As parenting practices have not been previously evaluated for family interventions such as FFT and MST direct comparisons based on an adolescent population are not available.

Parent-child bonding. There is a significant association between parent child attachment or bonding and subsequent psychological and conduct related problems including aggression (Raudino, Fergusson & Horward, 2013; Savage, 2014). Few studies report treatment related changes in bonding for conduct-problem interventions with adolescents. One exception is Slesnik & Prestopnik (2005) who describe a family based intervention for substance using adolescents; both family therapy and services as usual achieved medium effects pre-post on the parental care dimension of an attachment measure. Whilst the present study yielded only a small, non-significant effect for changes in parent-child bonding a regression analysis found a significant relationship between bonding and reduction in conduct problems. This suggests that the quality of the parent-child relationship is an important focus in FFT.

Parent-child aggression and neglect. Parental verbal and physical aggression to the child is both a sequela of childhood conduct problems amongst parents and a correlate of child conduct problems amongst children (Kolko, Dorn, Bukstein & Burke, 2008; Raudino et al., 2012). Interventions to reduce parental mistreatment have had variable success; Dubowitz, Feigelman, Lane & Kim (2009) reported small effect sizes for verbal and physical aggression in a primary care prevention study however Sturrock & Grey (2013) in an Incredible Years study reported medium effects ($d = 0.46-0.60$) for these factors. The present study yielded essentially no change in neglect and physical aggression ($d = -0.06$ to 0.07) and a small

positive effect upon verbal aggression ($d = 0.30$). The low level of severe problems at the outset of the study may have limited the scope for improvement and thus effect size; 85% of parents reported nil physical aggression and 86% reported nil neglect at pre-test. By way of comparison only 15% of parents reported nil verbal aggression at pre-test. In order to examine this issue effects were calculated for those parents reporting at least some physical aggression and neglect at pre-test. Cohen's d was estimated as 1.18 (parent aggression), 1.30 (parent neglect) and 1.70 (partner aggression)²⁶ all large effects. These figures suggest substantial improvements in parent and partner physical aggression and neglect among families where this occurs. However this must be interpreted with caution due to the very small numbers of participants involved (4-5).

Parent conflict. Young people who experience high levels of intimate partner violence between their parents are at risk of a range of adverse outcomes during childhood and young adulthood including conduct problems as a child and substance use, mental health issues and inter-partner violence as young adults (Fergusson & Horwood, 1998; Fergusson, Boden & Horwood, 2006). Interventions for IPV in the literature have included specific models designed for this problem (e.g. Wray, Hoyt & Gerstle, 2013). Some studies report only modest success (Babcock, Green & Robie, 2004). In the present study medium effects were obtained for reductions in the primary caregiver's violence to a partner and improvements in relationship quality. There was minimal change in violence perpetrated by the partner and the effect size was close to zero. In contrast to parent-child aggression 87% of participants with partners reported at least some partner violence at pre-test and even when those reporting nil violence were excluded the effect size for violence reduction was near-nil indicating minimal change in the level of partner violence (Cohen's $d = 0.04$). Having said that, 25% of participants reported physical violence or threats with a weapon hence the violence in question was predominantly verbal, - swearing, shouting, threatening. Whilst the intimate-partner violence reported was largely mild it is of some concern that little impact was made on partner-perpetrated violence by FFT. One possible reason for this may be that the primary caregiver's partner may be less involved in the FFT process and may in some cases not attend sessions (i.e. is at work), hence the disparity in outcomes across parents.

Parent relationship quality. Parental relationship quality is one of a range of inter-related socio-environmental factors which is both a correlate of child conduct problems (e.g. Bornovalova, Blazei, Malone, McGue & Iacono, 2013) and an outcome in later life of social and family adversities (Fergusson, McLeod & Horwood, 2014). There is little research on changes in relationship quality in the context of interventions for conduct problems, however two studies which did examine the influence of the Triple P parenting programme on relationship quality both recorded medium effect sizes (Roux, Sofranoff & Sanders, 2013; Salari, Ralph & Sanders, 2014; $d=0.40, 0.45$) which are similar to those obtained in the present study ($d=0.45$). An improvement in relationship quality is unsurprising given the centrality of relationship dynamics to the FFT model.

Delinquent peers. Association with delinquent peers is one of the most significant risk factors for conduct problems in adolescence (Dishion et al., 1999; Ronan & Curtis, 2008).

²⁶ Partner neglect could not be calculated; only one case of partner neglect at pre-test was reported.

The process whereby youth are introduced to and encouraged to partake of delinquent activities has been termed "deviancy training" and minimisation of contact with deviant youth has become a core treatment goal in a number of evidence-based interventions (Leve & Chamberlain, 2005; Huey, Henggeler, Brondino & Pickrel, 2000). Consequently it is of some concern that a parent-rated measure of the target child's association with delinquent peers evidenced almost no change over the course of the study (cf. Chapter 7, Table 10). Examination of data from the pre-test interview showed that approximately 30-50% of parents were reporting nil or minimal delinquent peer association prior to commencement of FFT however a similar proportion of children were reported to have anti-social friends and to be influenced by these friends. Interestingly when the young persons themselves were questioned about the level of delinquency amongst their friends they reported a significant reduction. As much youth delinquent activity is covert and unknown to parents the present results are consistent with parental concerns lagging behind an actual reduction in association with delinquent peers as reported by the youth.

Maternal depression. There is an association between a history of maternal depression and increased levels of child conduct problems (Fergusson & Lynsky, 1993; Kim-Cohen et al., 2005). As family and social dysfunction may be related to maternal depression a family therapy intervention may be effective for the alleviation of maternal depression and thus have an indirect influence on child behaviour via improved parental mood. This seems to be the case in the present study; the percentage of caregivers reporting a significant level of depression reduced from 34% at the beginning of the study to 19% at the end. This change displayed a trend towards significance ($p=0.10$). Further, there was a reduction in the severity of the depression symptoms reported by those reporting ongoing depression, to a degree consistent with a small to medium effect size, although this was not significant.

Treatment fidelity. In general no relationship was found between conduct problem outcomes and ratings of therapist treatment fidelity. This may be a measurement or statistical issue; the ratings may be inaccurate, they may be too uniform or the influence of fidelity may be masked by other stronger effects. Treatment fidelity has been found to be a significant factor in previous FFT research; Sexton & Turner (2011) in a large FFT study obtained essentially a nil effect overall however when data for those therapists rated as displaying good fidelity was analysed separately the outcomes for this group were significant. Clearly the poorer outcomes which impaired the study as a whole were being obtained by therapists who were not following the FFT model to an optimum degree. In the present study fidelity ratings were generally high; for instance the overall average ratings for dissemination fidelity and overall fidelity across all participants and therapists were 5.26 and 4.42 respectively in a 6 point scale. Thus one possible explanation is that variation in fidelity had little influence on treatment as fidelity was at an optimum level.

12.6 Limitations

There are three key limitations to this study:

Firstly, in the absence of a control group it cannot be stated with absolute certainty that the positive changes in conduct problems and other outcomes are due primarily to the FFT intervention. Plausible alternative sources of an apparent treatment effect include maturation, regression to the mean and non-specific treatment effects²⁷.

Secondly, it was not possible to collect official offending data. Whilst official data is not without its weaknesses, for instance it tends to under-report all but the more severe anti-social behaviours, official data does connect the intervention with a quantifiable socially and politically important outcome. The use of offending data would also permit a more thorough comparison with the existing FFT, MST and delinquency prevention literatures.

Thirdly, the present study experienced a significant attrition of participants both drop-outs from the FFT treatment and those lost to the research whether or not they completed treatment. Unlike the previous limitations participant attrition is potentially remediable and may have affected outcomes and therefore bears some discussion.

Nineteen percent of those contacted about the study declined to participate at the outset. A matter of some consideration was that the ethical guidelines for this study did not allow financial compensation for participation in the parent research interview, which was of significant duration (1 to 2 hours). Families were referred to the FFT programme by Child Youth & Family at the initiative of Child Youth & Family and whilst treatment was not compulsory family attitudes to the opportunity of treatment were varied. Of the 59 parents or caregivers who agreed to participate and completed the pre-test 29% (17) missed one or other of the post or follow-up parent interviews. Of those missing an assessment 70% (12) had terminated treatment early according to the FFT therapist. The most common rationale for treatment termination was family avoidance or withdrawal from treatment. Thus seemingly families unmotivated to complete further treatment sessions were also unmotivated to complete further interviews. Thus whilst the rate of attrition was not salutary it was most commonly related to treatment drop-out. The attrition rate amongst those families who completed treatment was 12% (5 of 40).

Treatment drop-out is a significant issue in the treatment of conduct problems with drop-out rates of 50%-75% given by some authors (Kazdin, 1990). In a study of community treatments for disruptive behaviour in Greece 46% of cases did not complete treatment (Anagnostopoulos et al., 2006). Where cited, treatment drop-out rates in the FFT literature range from 14% (Waldron, 2001) to 23% (White, 2013). Data loss rates may exceed drop-out rates where some participants who complete treatment nonetheless miss assessments, - the data loss rate in the latter study (missing at least one measure) was 43%. In the MST literature treatment drop-out and data loss rates tend to be low however in most major

²⁷ Maturation refers to developmental changes in young persons which may lead to improved functioning. Regression to the mean is the statistical tendency for extreme scores on a measure to become less extreme when measured repeatedly. Non-specific effects are general aspects of all therapies such as therapist skill or personaility or the effect of professional attention rather than any specific methodology.

studies participants are paid to attend assessment sessions (e.g. Henggeler et al., 1997). In an earlier study where financial incentives were not used the attrition rate was 23% (Henggeler Melton & Smith, 1992).

The primary concern with participant attrition in the conduct problem literature is the threat to internal validity of mortality bias; that is, the differential loss to the study of poorly performing participants thus leading to an inflated degree of improvement amongst those remaining in the study (Kazdin, 1990). As noted in Chapter 3 participants with missing data obtained somewhat larger (more severe) scores at pre-test in three of seven measures suggestive of a relationship between symptom severity and treatment drop-out or assessment compliance problems. However as shown above, for those participants remaining in the study, those with higher than average scores at pre-test tended to obtain larger effect sizes, thus the direction of any mortality bias may be towards diminishing rather than exaggerating the effect sizes obtained. Whilst any bias is regrettable, the evidence relating to mortality bias is mixed.

12.7 Future directions

- Given that the present study lacked a control group, a future study undertaken as a randomised controlled trial would add significantly to the robustness of the findings.
- Whether or not a randomised controlled trial is feasible replication of the present study across different FFT sites will be important for the generalization of study findings.
- A series of small studies using a within-subject design would provide useful additional data and may be logistically easier to achieve than larger studies. For instance a multiple-baseline design would involve collecting daily or weekly behavioural measures before and after the commencement of FFT and tracking changes over time.
- Further investigation is required to understand those aspects of FFT associated with different outcomes for Māori, with a view to maximising effectiveness. For instance, does the ethnicity of the therapist affect outcome for Māori? If the study is large enough statistical techniques may be used to identify the predictors of better outcomes for Māori.
- Further studies would benefit from a careful examination of the potential influence of therapist treatment fidelity. It may be beneficial to specifically test the reliability of therapist ratings by having more than one expert fidelity rater and to use more exacting measurement techniques such as observations of therapy sessions.
- A focus on the minimisation of attrition will be important in future research with at the very least modest financial compensation being provided for interview completion. Other key strategies may include the creation of a "hard to contact"

team tasked solely with securing interviews with participants at risk of attrition, securing multiple alternative contact details from families at the first interview, closer cooperation with clinicians and establishing certain measures as an integral component of the clinical programme, as opposed to research “add ins”.

- Future studies would also benefit from using official crime data as an additional measure to verify the young person’s progress. Whilst official data has weaknesses, for instance much crime is unreported, it has the advantage of being reliably obtainable for many years after the intervention and it is the established rubric for crime outcomes at the legal and political level.
- Future study designs could include follow-up assessments one or more years post-treatment to ascertain the degree to which treatment gains are maintained in the longer term.

APPENDIX 1. New Zealand MST replications.

All effect sizes are Cohen's d.

Curtis, Ronan, Heiblum & Crellin (2009)

Independent replication, n=65. Offending frequency effect size 0.23, severity 0.16. MST Behaviour rating scale (youth compliance, family communication, family relations) 0.29. All effect sizes pre-post.

Russell (2008)

Unpublished PhD study, independent replication, n=73. Offending frequency 0.13, Severity 0.07. "Negative behaviour" (internalising and externalising) 0.67. All effect sizes pre-post.

Grace, McLean & Warren (2006)

Evaluation of Reducing Youth Offending Programme, n=variable 17-172²⁸. Conduct Disorder rating scale -0.08, Aggression rating scale 0.37, Attention problems 0.35, Self reported delinquency 0.29, Offence frequency 0.41, Offence severity 0.53. All effect sizes pre-post.

²⁸ The core of this study is a complex re-offending model which is not comparable with any of the research referred to in this study, however some selected outcomes are reported in enough detail to allow calculation of effects. Offense data is originally from MST New Zealand.

References

- Alexander, J. J., & Parsons, B. F. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81*(3), 219-225.
- Anagnostopoulos, D. C., Vlassopoulos, M., Lazaratou, H., Tzavara, C., Zelios, G., & Ploumpidis, D. (2006). Evaluating mental health services in a Greek community: The factor of non-compliance to therapy. *European Child & Adolescent Psychiatry, 15*(8), 435-441.
- Armsden, G. C., & Greenberg, M. T. (1987). The Inventory of Parent and Peer Attachment: Individual Difference and Their Relationship to Psychological Well-Being in Adolescence. *Journal of Youth and Adolescence, 16*(5), 427-454.
- Arnold, P., O-Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment, 5*(2), 137-144.
- Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review, 23*(8), 1023-1053.
- Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of multisystemic therapy and functional family therapy effectiveness: A multiyear statewide propensity score matching analysis of juvenile offenders. *Criminal Justice and Behavior, 41*(9), 1033-1056.
- Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of multisystemic therapy and functional family therapy effectiveness: A multiyear statewide propensity score matching analysis of juvenile offenders. *Criminal Justice and Behavior, 41*(9), 1033-1056.
- Baldwin, S. A., Christian, S., Berkeljon, A., Shadish, W. R., & Bean, R. (2012). The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital & Family Therapy, 38*(1), 281-304.
- Barton, C., Alexander, J. F., Waldron, H., Turner, C. W., & Warburton, J. (1985). Generalizing treatment effects of functional family therapy: three replications. *The American Journal of Family Therapy, 13*(3), 16-26.
- Barton, C., Alexander, J. F., Waldron, H., Turner, C. W., & Warburton, J. (1985). Generalizing treatment effects of Functional Family Therapy: Three replications. *American Journal of Family Therapy, 13*(3), 16-26.
- Beauchaine, T. P., Webster-Stratton, C., & Reid, M. J. (2005). Mediators, Moderators, and Predictors of 1-Year Outcomes Among Children Treated for Early-Onset Conduct Problems: A Latent Growth Curve Analysis. *Journal of Consulting and Clinical Psychology, 73*(3), 371-388.
- Bjorknes, R., Kjobli, J., Manger, T., & Jakobsen, R. (2012). Parent training among ethnic minorities: Parenting practices as mediators of change in child conduct problems. *Family Relations: An Interdisciplinary Journal of Applied Family Studies, 61*(1), 101-114.
- Bornovalova, M. A., Blazei, R., Malone, S. H., McGue, M., & Iacono, W. G. (2013). Disentangling the relative contribution of parental antisociality and family discord to child disruptive disorders. *Personality Disorders: Theory, Research, and Treatment, 4*(3), 239-246.

- Braiker, H., & Kelley, H. H. (1979). Conflict in the development of close relationships. In R. L. Burgess & T. L. Huston (Eds.), *Social exchange in developing relationships*. New York: Academic.
- Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004). Multisystemic Treatment: A Meta-Analysis of Outcome Studies. *Journal of Family Psychology*, 18(3), 411-419.
- Curtis, N. M., Ronan, K. R., Heiblum, N., & Crellin, K. (2009). Dissemination and Effectiveness of Multisystemic Treatment in New Zealand: A Benchmarking Study. *Journal of Family Psychology*, 23(2), 119-129.
- Dadds, M., Majeaun, A., & Fraser, J. A. (2003). Parenting and Conduct Problems in Children: Australian Data and Psychometric Properties of the Alabama Parenting Questionnaire. *Australian Psychologist*, 38(4), 238-241.
- Dadds, M., & Powell, M. (1991). The relationship of interparental conflict and global marital adjustment to aggression, anxiety and immaturity in aggressive and non-clinic children. *Journal of Abnormal Child Psychology*, 19, 553-561.
- Datchi, C. C., & Sexton, T. L. (2013). Can family therapy have an effect on adult criminal conduct? Initial evaluation of functional family therapy. *Couple and Family Psychology: Research and Practice*, 2(4), 278-293.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, 54(9), 755-764.
- Dishion, T. J., & Patterson, G. R. (1992). Age Effects in Parent Training Outcome. *Behavior Therapy*, 23, 719-729.
- Dubowitz, H., Feigelman, S., Lane, W., & Kim, J. (2009). Pediatric primary care to help prevent child maltreatment: The Safe Environment for Every Kid (SEEK) Model. *Pediatrics*, 123(3), 858-864.
- Duncan, T. M., Davey, M., & Davey, A. (2011). Transporting functional family therapy to community-based programs. *The Family Journal*, 19(1), 41-46.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-Analysis of Experiments With Matched Groups or Repeated Measures Designs. *Psychological Methods*, 1(2), 170-177.
- Elliott, D. S. (Ed.). (2000). *Blueprints for Violence Prevention. Book Three Functional Family Therapy*. Golden, Colorado: Venture Publishing.
- Fabiano, G. A., Pelham, W. E., Jr., Coles, E. K., Gnagy, E. M., Chronis-Tuscano, A., & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, 29(2), 129-140.
- Farrington, D. P., & Welsh, B. C. (2003). Family-based prevention of offending: A meta-analysis. *Australian and New Zealand Journal of Criminology*, 36(2), 127-151.
- Ferguson, D. M., & Horwood, L. J. (1998). Exposure to interparental violence in childhood and psychological adjustment in young adulthood. *Child Abuse & Neglect*, 22(5), 339-357.
- Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2006). Examining the intergenerational transmission of violence in a New Zealand birth cohort. *Child Abuse & Neglect*, 30, 89-108.
- Fergusson, D. M., & Horwood, L. J. (1996). The Role Of Adolescent Peer Affiliations in the Continuity Between Childhood Behavioral Adjustment and Juvenile Offending. *Journal of Abnormal Child Psychology*, 24, 205-221.

- Fergusson, D. M., & Horwood, L. J. (1999). Prospective Childhood Predictors of Deviant Peer Affiliations in Adolescence. *Journal of Child Psychology and Psychiatry*, 40(4), 581-592.
- Fergusson, D. M., McLeod, G. F., & Horwood, L. (2014). Parental separation/divorce in childhood and partnership outcomes at age 30. *Journal of Child Psychology and Psychiatry*, 55(4), 352-360.
- Fossum, S., Handegard, B. H., Martinussen, M., & Mørch, W. T. (2008). Psychosocial interventions for disruptive and aggressive behaviour in children and adolescents: A meta-analysis. *European Child & Adolescent Psychiatry*, 17(7), 438-451.
- Frick, P. J., Lahey, B. B., Loeber, R., Stouthamer-Loeber, M., Christ, M. A. G., & Hanson, K. (1992). Familial risk factors to oppositional defiant disorder and conduct disorder: Parental psychopathology and maternal parenting. *Journal of Consulting and Clinical Psychology*, 60(1), 49-55.
- Friedman, A. S. (1989). Family Therapy Vs. Parent Groups: Effects on Adolescent Drug Abusers. *The American Journal of Family Therapy*, 17(4), 335-347.
- Gordon, D. A. (1995). Functional Family Therapy for delinquents. In R. Ross, D. H. Antonowicz & G. K. Dhaliwari (Eds.), *Going straight: Effective delinquency prevention and offender rehabilitation* (pp. 163-178). Ontario: Air training and publications.
- Gordon, D. A., Arbuthnot, J., Gustafson, K. E., & Mcgreen, P. (1988). Home-based behavioral-systems family therapy with disadvantaged juvenile delinquents. *The American Journal of Family Therapy*, 16(3), 243-255.
- Gordon, D. A., Graves, K., & Arbuthnot, J. (1995). The effect of functional family therapy for delinquents on adult criminal behavior. *Criminal Justice and Behavior*, 22(1), 60-73.
- Gordon, D. A., Graves, K., & Arbuthnot, J. (1995). The effect of functional family therapy for delinquents on adult criminal behaviour. *Criminal Justice and Behaviour*, 22(1), 60-73.
- Grace, R., McLean, A., & Warren, J. (2006). *An Evaluation of the Reducing Youth Offending Programme*. Departments of Child, Youth and Family and Corrections
- Graham, C., Carr, A., Rooney, B., Sexton, T., & Satterfield, L. R. W. (2014). Evaluation of functional family therapy in an Irish context. *Journal of Family Therapy*, 36(1), 20-38.
- Henggeler, S. W., & Ashli, J. (2012). Empirically supported family-based treatments for conduct disorder. *Journal of Marital & Family Therapy*, 38(1), 30-58.
- Henggeler, S. W., Clingempeel, W., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41(7), 868-874.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology*, 65(5), 821-833.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, 60(6), 953-961.

- Huey, S. J., Jr., Henggeler, S. W., Brondino, M. J., & Pickrel, S. G. (2000). Mechanisms of change in multisystemic therapy: Reducing delinquent behavior through therapist adherence and improved family and peer functioning. *Journal of Consulting and Clinical Psychology, 68*(3), 451-467.
- Kazdin, A. E. (1990). Premature termination from treatment among children referred for antisocial behavior. *Child Psychology & Psychiatry & Allied Disciplines, 31*(3), 415-425.
- Kolko, D. J., Dorn, L. D., Bukstein, O., & Burke, J. D. (2008). Clinically Referred ODD Children with or without CD and Healthy Controls: Comparisons Across Contextual Domains. *Journal of Child and Family Studies, 17*, 714-734.
- Lee, S., Aos, S., Drake, E., Pernucci, A., Miller, M., & Anderson, L. (2012). Return on Investment: Evidence-Based Options to Improve Statewide Outcomes, April 2012 (Document No. 12-04-1201) (pp. 1-8). Olympia: Washington State Institute for Public Policy.
- Leve, L. D., & Chamberlain, P. (2005). Association with Delinquent Peers: Intervention Effects for Youth in the Juvenile Justice System. *Journal of Abnormal Child Psychology, 33*(3), 339-347.
- Marie, D., Fergusson, D., & Boden, J. (2009). Ethnic Identity and Exposure to Maltreatment in Childhood: Evidence from a New Zealand Birth Cohort. *Social Policy Journal of New Zealand, 36*, 154-171.
- Marie, D., Fergusson, D. M., & Boden, J. M. (2008). Ethnic identification, social disadvantage, and mental health in adolescence/young adulthood: results of a 25 year longitudinal study. *Australian and New Zealand Journal of Psychiatry, 42*, 293-300.
- Marie, D., Fergusson, D. M., & Boden, J. M. (2008). Links between ethnic identification, cannabis use and dependence, and life outcomes in a New Zealand birth cohort. *Australian and New Zealand Journal of Psychiatry, 42*, 780-788.
- Marie, D., Fergusson, D. M., & Boden, J. M. (2009). Ethnic identity and criminal offending in a New Zealand birth cohort. *Australian and New Zealand Journal of Criminology, 42*(3), 354-368.
- McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential Effectiveness of Behavioral Parent-Training and Cognitive-Behavioral Therapy for Antisocial Youth: A Meta-Analysis. *Journal of Abnormal Child Psychology, 34*(4), 527-543.
- Menting, A. T., Orobio de Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behaviour: A meta-analytic review. *Clinical Psychology Review, 33*, 901-913.
- Ogden, T., & Hagen, A. (2006). Multisystemic Treatment of Serious Behaviour Problems in Youth: Sustainability of Effectiveness Two Years after Intake. *Child and Adolescent Mental Health, 11*(3), 142-149.
- Ogden, T., & Halliday-Boykins, C. A. (2004). Multisystemic Treatment of Antisocial Adolescents in Norway: Replication of Clinical Outcomes Outside of the US. *Child and Adolescent Mental Health, 9*(2), 77-83.
- Ozechowski, T. J. (2014). Empirical Bayes MCMC estimation for modeling treatment processes, mechanisms of change, and clinical outcomes in small samples. *Journal of Consulting and Clinical Psychology, 82*(5), 854-867.
- Raudino, A., Fergusson, D. M., & Horwood, L. J. (2013). The quality of parent/child relationships in adolescence is associated with poor adult psychosocial adjustment. *Journal of Adolescence, 36*, 331-340.

- Rhoades, K. A., & O-Leary, S. G. (2007). Factor Structure and Validity of the Parenting Scale. *Journal of Clinical Child and Adolescent Psychology*, 36(2), 137-146.
- Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid depressive and substance use disorders. *Journal of Consulting and Clinical Psychology*, 82(2), 342-348.
- Ronan, K. R., & Curtis, N. (2008). Treatment of antisocial youth and families *Innovations in clinical practice: Focus on group, couples, & family therapy* (pp. 5-27).
- Rowland, M. D., Halliday-Boykins, C. A., Henggeller, S. W., Cunningham, P. B., Lee, T. G., Kruesi, M. J., et al. (2005). A Randomized Trial of Multisystemic Therapy With Hawaii's Felix Class Youth. *Journal of Emotional and Behavioral Disorders*, 13(1), 13-23.
- Russell, C. J. (2008). *Multisystemic therapy in New Zealand: Effectiveness and predictors of outcome*. (Unpublished doctoral dissertation), Massey University, Palmerston North.
- Salari, R., Ralph, A., & Sanders, R. (2014). An Efficacy Trial: Positive Parenting Program for Parents of Teenagers. *Behaviour Change*, 31(1), 34-52.
- Sanders, M. R., Baker, S., & Turner, K. M. (2012). A randomized controlled trial evaluating the efficacy of Triple P Online with parents of children with early-onset conduct problems. *Behaviour Research and Therapy*, 50, 675-684.
- Savage, J. (2014). The association between attachment, parental bonds and physically aggressive and violent behaviour: A comprehensive review. *Aggression and Violent Behaviour*, 19, 164-178.
- Sexton, T., & Turner, C. W. (2010). The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. *Journal of Family Psychology*, 24(3), 339-348.
- Sexton, T., & Turner, C. W. (2011). The Effectiveness of Functional Family Therapy for Youth with Behavioral Problems in a Community Practice Setting. *Couple and Family Psychology: Research and Practice*, 1(Spring), 3-15.
- Sexton, T., & Turner, C. W. (2011). The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. *Couple and Family Psychology: Research and Practice*, 1(S), 3-15.
- Shaffer, A., Lindhiem, O., & Kolko, D. J. (2013). Treatment effects of a modular intervention for early-onset child behavior problems on family contextual outcomes. *Journal of Emotional and Behavioral Disorders*, 21(4), 277-288.
- Shelton, K. K., Frick, P. J., & Wootton, J. (1996). Assessment of Parenting Practices in Families of Elementary School-Age Children. *Journal of Clinical Child Psychology*, 25(3), 317-329.
- Slesnick, K., & Frick, P. J. (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology*, 25, 317-329.
- Slesnick, N., & Prestopnik, J. L. (2004). Office versus Home-Based Family Therapy for Runaway, Alcohol Abusing Adolescents: Examination of Factors Associated with Treatment Attendance. *Alcoholism Treatment Quarterly*, 22(2), 3-19.
- Slesnick, N., & Prestopnik, J. L. (2005). Ecologically based family therapy outcome with substance abusing runaway adolescents. *Journal of Adolescence*, 28(2), 277-298.

- Slesnick, N., & Prestopnik, J. L. (2009). Comparison of Family Therapy Outcome With Alcohol-Abusing, Runaway Adolescents. *Journal of Marital and Family Therapy*, 35, 255-277.
- Slesnick, N., & Prestopnik, J. L. (2009). Comparison of family therapy outcome with alcohol-abusing, runaway adolescents. *Journal of Marital and Family Therapy*, 35(3), 255-277.
- Straus, M. A. (1979). Measuring Interfamily Conflict and Violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family*, 41(1), 75-88.
- Straus, M. A., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child-maltreatment with Parent-Child Conflict Tactic Scales: Development and psychometric data for a national sample of American parents. *Child Abuse and Neglect*, 22(4), 249-270.
- Sturrock, F., & Gray, D. (2013). *Incredible Years Pilot Study: Evaluation Report*. Wellington: Ministry of Social Development NZ.
- Sundell, K., Hansson, K., Lofholm, C. A., Olsson, T., Gustle, L., & Kadesjo, C. (2008). The Transportability of Multisystemic Therapy to Sweden: Short-Term Results From a Randomized Trial of Conduct-Disordered Youth. *Journal of Family Psychology*, 22(3), 550-560.
- van der Put, C. E., Asscher, J. J., Stams, G. J. J., van der Laan, P. H., Breuk, R., Jongman, E., et al. (2013). Recidivism after treatment in a forensic youth-psychiatric setting: The effect of treatment characteristics. *International Journal of Offender Therapy and Comparative Criminology*, 57(9), 1120-1139.
- Waldron, H. B., Slesnick, N., Brody, J. L., Turner, C. W., & Peterson, T. R. (2001). Treatment Outcomes for Adolescent Substance Abuse at 4- and 7-Month Assessments. *Journal of Consulting and Clinical Psychology*, 69(5), 802-813.
- Webster-Stratton, C., Reid, M. J., & Beauchaine, T. P. (2013). One-Year Follow-Up of Combined Parent and Child Intervention for Young Children with ADHD. *Journal of Clinical Child & Adolescent Psychology*, 42(2), 251-261.
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating Children With Early-Onset Conduct Problems: Intervention Outcomes for Parent, Child, and Teacher Training. *Journal of Clinical Child and Adolescent Psychology*, 33, 105-124.
- White, S. F., Frick, P. J., Lawing, K., & Bauer, D. (2013). Callous-unemotional traits and response to functional family therapy in adolescent offenders. *Behavioral Sciences and the Law*, 31, 271-285.
- White, S. F., Frick, P. J., Lawing, K., & Bauer, D. (2013). Callous-unemotional traits and response to Functional Family Therapy in adolescent offenders. *Behavioral Sciences & the Law*, 31(2), 271-285.
- Wilson, D. B., & Lipsey, M. W. (2001). The Role of Method in Treatment Effectiveness Research: Evidence from Meta-Analysis. *Psychological Methods*, 6(4), 413-429.
- Wray, A. M., Hoyt, T., & Gerstle, M. (2013). Preliminary Examination of a Mutual Intimate Partner Violence Intervention Among Treatment-Mandated Couples. *Journal of Family Psychology*, 27(4), 664-670.
- Zazzali, J. L., Sherbourne, C., Hoagwood, K. E., Greene, D., Bigley, M. F., & Sexton, T. L. (2008). The adoption and implementation of an evidence based practice in child and family mental health services organizations: A pilot study of functional family therapy in New York State. *Administration and Policy in Mental Health and Mental Health Services Research*, 35(1-2), 38-49.

