

A pilot study of functional family therapy in New Zealand

Charles Heywood Youth Horizons Trust, David Fergusson University of Otago, New Zealand

Conduct problems affect between five and ten percent of children in New Zealand and are associated with a wide-range of adverse outcomes in later life. There is a clear need for the verification of evidence-based interventions for the New Zealand population including Māori. A pilot study of Functional Family Therapy (FFT) found significant ($p < 0.01$) pre to follow-up reductions in two of five conduct problem measures and medium effect sizes for four of five measures. Outcomes for Māori and non-Māori were similar. Variations in Therapist treatment fidelity and competence did not significantly influence outcomes. Parent satisfaction with FFT was high and Māori parents' satisfaction with cultural aspects of FFT very high. These findings provide initial evidence that FFT is effective in the New Zealand context.

Keywords: Functional family therapy, conduct problems, treatment effectiveness, transportability, outcomes for Māori

Conduct problems including Conduct Disorder and Oppositional Defiant Disorder affect 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, reduced life expectancy, domestic violence, poor educational achievement, unemployment, welfare dependency and conflict with parents and teachers (Fergusson, Horwood, & Ridder, 2005; Kratzer & Hodgins, 1997; Offord & Bennett, 1994). Interventions for conduct problems have been well researched in the international literature and a number of programmes have been identified as effective in authoritative reviews (Eyberg, Nelson, & Boggs, 2008). Although some research has been conducted with respect to interventions for younger children (Fergusson, Horwood & Stanley 2013, Fergusson, Horwood & Stanley 2009) there has been very little research published in New Zealand with respect to interventions for older children and adolescents with conduct problems; a study of Multi-systemic Therapy, Curtis, Ronan, Heiblum and Crellin (2009), appears to be the only example. Of concern in the

New Zealand context is the high rate of conduct problems identified amongst Māori youth with rates of conduct disorder ranging from 1.9 to 4.5 times that of the non-Māori population as ascertained in two large longitudinal studies (Fergusson, Poulton, Horwood, Milne and Swain-Campbell, 2003). As one investigation of Māori youth offending notes "On average Māori youth are three times more likely to be apprehended, prosecuted and convicted than non-Māori youth" (Owen, 2001, cited in Curtis, Ronan, Heiblum, Reid & Harris, 2002). Although comprising 15% of the population in the 2013 census as of 2016 Māori comprise 51% of the prison population (Statistics New Zealand, 2016). These proportions are mirrored by the high rate of involvement of Māori youth and families in programmes for the treatment of conduct problems; within Youth Horizons, New Zealand's largest provider of evidence based interventions for conduct problems, 45% of children participating in treatment were Māori in 2015 (Youth Horizons, 2015).

A number of interventions have been systematically evaluated for the treatment of child and adolescent conduct problems in recent decades. Family interventions have shown particular promise and one such intervention is Functional Family Therapy (FFT). 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 FFT is now implemented in 220 sites across five countries and claims to work with more families per year than any other evidence based intervention (fftlc.com). FFT is also recognised as cost effective with better than average cost-benefit ratios when compared to other juvenile justice interventions (Washington state institute for public policy, 2016).

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. FFT also uses cognitive-behavioural techniques such as reframing and emotion-management strategies.

FFT is a short-term intervention of 8-12 sessions over a period of 2-4 months. 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 has been appraised favourably in a number of reviews (e.g. Baldwin, Christian, Berkeljon, Shadish & Bean, 2012; Carr, 2014; Henggeler & Sheidow, 2012) however its performance has been superior in the context of high-quality efficacy trials than in community based effectiveness studies. An initial series of University-based studies with juvenile offenders established the efficacy of FFT under ideal conditions with reductions in recidivism ranging from 35% to 84% relative to an alternative treatment (Alexander & Parsons, 1973; Barton, Alexander, Waldron, Turner & Warburton, 1985; Gordon, Arbuthnot, Gustafson & McGreen, 1988; Klein, Alexander & Parsons, 1977). Maintenance of treatment effects was demonstrated with follow-up assessments of up to five years post-treatment (Gordon, Graves & Arbuthnot, 1995). A single efficacy study, Waldron, Slesnick, Brody, Turner & Peterson (2001) failed to find a significant reduction in marijuana use and externalising behaviour for FFT relative to Cognitive Behaviour Therapy and Group Therapy control groups.

Community-based effectiveness trials have obtained variable outcomes. Of six published effectiveness studies conducted in the United States, four report favourable outcomes: Lantz (1982), as described in Elliot, Alexander, Pugh, Parsons and Sexton (2000), obtained significant reductions in recidivism relative to an alternative treatment. Rohde, Waldron, Turner, Brody and Jorgensen (2014) found FFT followed or preceded by cognitive behaviour therapy (CBT) to result in greater reductions in substance use than a combined FFT/CBT intervention. Sleznick and Prestopnik (2009) and White, Frick, Lawing and Bauer (2013) found significant pre-post reductions in drug and alcohol use and conduct problems respectively. However two studies failed to obtain statistically significant results; Friedman (1989) found that FFT did not reduce drug use or externalising behaviour relative to a parent group intervention and Sexton & Turner (2010) in a large randomised controlled trial (n=917) found identical recidivism rates for FFT and probation services as usual 12 months post-treatment. In a post-hoc analysis the

authors of the latter study found that recidivism was related to treatment fidelity and therapists high in adherence to the FFT treatment model did indeed achieve significant reductions in felony and violent crimes relative to control.

Two transportability studies conducted in community settings in Ireland have demonstrated significant reductions in adolescent conduct problems; Graham, Carr, Rooney, Sexton and Sattersfield (2014) demonstrated improvements relative to baseline and in a randomised controlled trial Hartnett, Carr & Sexton (2015) found improvements in conduct problems and family adjustment relative to a wait-list control. Graham et al. (2014) also found that treatment outcome was mediated by treatment fidelity with those therapists demonstrating better fidelity obtaining better outcomes. In addition two studies conducted in Sweden, although not published in English, are reported in secondary sources to have found lower rates of recidivism in FFT as compared to services as usual (Hansson, Johansson, Drott-Englen & Benderix 2004 as reported in Robbins 2016; Hansson 1998, as reported in Elliott et al., 2000).

Thus whilst FFT is an intervention which can achieve strong outcomes when well implemented, a less than optimum - or merely different - implementation reflective of local conditions, client populations or workforce may result in less than optimum outcomes. In this context it has been argued that "empirically supported" interventions validated with primarily European participants cannot be said to be generalizable to diverse cultural or ethnic groups unless specifically tested with these groups (Cardemil, 2010; La Roche & Christopher, 2008). As the majority of families participating in FFT studies to date have been Caucasian (cf. Henggeler & Sheidow, 2012), and given the risks posed by untreated or ineffectively treated conduct problems for youth in general and Māori in particular, FFT's effectiveness and acceptability in New Zealand requires verification.

This paper describes a pilot study of FFT as delivered by a community organisation in Auckland, New Zealand. The study was intended to test five hypotheses: Firstly, that FFT as delivered in a community setting in New Zealand

will obtain significant reductions in conduct problems comparable to those achieved previously in effectiveness and transportability studies. Secondly, conduct problem outcomes for Māori will be similar in magnitude to those for non-Māori. Thirdly, parents will express a high degree of satisfaction with the FFT intervention. Fourthly, Maori parents will express a high degree of satisfaction with culturally relevant aspects of the FFT intervention. Fifthly, participants treated with a higher level of fidelity and competence will experience better outcomes.

Method

Design

The study was conducted at Youth Horizons, a community based non-government organisation in Auckland, New Zealand. FFT had been introduced to New Zealand by Youth Horizons in 2009. Whilst the intervention was not specifically adapted for Māori it was delivered in a culturally informed manner which respected common protocols involved in working therapeutically with Māori in their own home. The design was a single-group outcome study with assessment points at the commencement of treatment (pre), six-months later (post) and twelve months post treatment commencement (follow-up). All participants who received at least one treatment session were deemed to have started FFT and were to be assessed at these times whether or not they completed treatment.

Participants

Participants in the study were 59 young persons and families referred to Youth Horizons FFT by child welfare services (Child Youth & Family) between January 2011 and October 2012. Consent for study participation was sought from all FFT referrals during this period except where the family had been assigned to an FFT therapist with less than 3 months experience, or the family was unable to be contacted by the researcher prior to treatment commencement. Nineteen percent of families contacted declined to participate in the research.

Inclusion criteria were (a) child conduct problems as the primary referral

concern, (b) the child must be aged between 9 and 16 years¹, and (c) the young person must be living in the family home with at least one parent or permanent caregiver. Exclusion criteria were (a) young persons with an intellectual disability, psychosis or pervasive developmental disorder and (b) young persons whose primary referral issue was sex-offending or substance use. Exclusions were made on the basis of information accompanying the referral document. Diagnoses of Conduct Disorder or Oppositional Defiant Disorder were not required for FFT or study inclusion. The age range of the youth was 9 to 16 years with a mean of 13 years 7 months. Boys outnumbered girls two to one (70% vs. 30% of the sample). Māori were the largest ethnic group (45%) followed by New Zealand European (33%), Cook Island Māori (10%), other European (7%), Tongan (3%), Niuean (2%) and Fijian (2%).

The families participating in this study evidenced a number of indices of social and economic disadvantage. For instance, more than three-quarters (78%) of the primary caregivers had at most a secondary school qualification. The most common source of income was a social welfare benefit (62%). A majority of families were solo-parent families (69%). The median weekly household income of participants (\$500) was approximately a third of the median household income in New Zealand at that time. Significant numbers of parents and caregivers reported problem issues amongst members of their extended family other than the child referred to FFT; problems with the Police (21% of families), drugs (14%), depression (33%), anxiety (40%) and suicide attempt (10%). Levels of physical punishment of the child were moderate (20% of parents) and 19% percent of parents with a current partner reported at least one instance of physical assault upon themselves in the previous six-months. Parents reported contact with multiple government agencies with 46% having contact with Child & Adolescent Mental Health Services and 23% having contact with Special Education services re child learning or behaviour within the last five years. Over a third of children

(39%) were not attending school at the time the study commenced. These findings are consistent with the literature relating social and economic factors to the development of conduct problems (Deater-Deckard, Dodge, Bates & Pettit, 1998; Fergusson et al., 2005).

Ethics

Ethics approval was sought and granted by the Northern X Regional Ethics Committee, reference number NTX/10/06/052. Informed consent was required from both parents/caregivers and young persons.

Therapists

Ten therapists were involved in this study. Each had received the FFT introductory training as delivered by instructors from FFT LLC who travelled from the United States to deliver the training. All therapists received ongoing supervision from the programme supervisor, who was in turn supervised by an experienced FFT clinician via skype and telephone from the United States. The therapists had from less than one year to three years experience of FFT. Six therapists were female, four male. Prior professional roles included Psychologist (5), Psychotherapist (2), intern-psychologist (1), Multisystemic therapy therapist (1) and youth personal advisor (1). All therapists were university graduates. Two therapists were of Māori descent.

Measures

Parent and Teacher rated Conduct Problems

Parallel parent and teacher measures of conduct problems were constructed based on the fifteen and eight DSM-IV criteria for Conduct Disorder (CD) and Oppositional Defiant Disorder (ODD) respectively. DSM-IV Conduct Disorder involves behaviours relating to physical aggression (e.g. initiates physical fights), deceitfulness and theft (e.g. has stolen while confronting a victim), property damage (e.g. has deliberately destroyed others' property) and violation of rules (e.g. is often truant). DSM-IV Oppositional Defiant Disorder involves less severe problem behaviours relating to negativity (e.g. often argues with others), hostility (e.g.

is often touchy and easily annoyed by others) and defiance (e.g. often actively defies or refuses to comply with adults' requests). The 23 items, each rated on a 1 to 3 point scale, were summed to create a total score representing the overall level of conduct problems. A measure with identical item content but using a four point scale, The Rating Scale for Disruptive Behaviour Disorders (Silva et al., 2005), has shown good to excellent internal consistencies as measured by coefficient alphas of between 0.78 to 0.96 across informants (teachers, parents) and constructs (Conduct Disorder, Oppositional Defiant Disorder). Alpha calculated for the present data (all 23 items) was 0.88, suggesting a high level of coherence within the combined CD/ODD measure.

Parent reported Young Person alcohol and substance use

An instrument was created whereby parents reported the young person's frequency of use, to the best of their knowledge, over the previous two months, of seven categories of substance ranging from alcohol and cannabis to amphetamines and opiates. For each category the parent reported the frequency of use as either daily, weekly, more than once a month, less than once a month, or not at all. Scores were pro-rated to a days-per-year equivalent within categories and aggregated across categories to give a total score.

Parent and Young Person reported delinquent behaviour

Parent and young person versions of an instrument based on the Self Report Delinquency Inventory (SRD: Elliott & Huizinga, 1989), were used to assess young person delinquent behaviours over the previous six months. Domains surveyed included property offences (burglary, theft, fire setting, property damage), violent offences (fighting, using a weapon, assault), use of alcohol and drugs (e.g. drinking alcohol during school hours) and other delinquent behaviours (struggling to get away from a policeman, running away from home, truancy). A count measure was derived reflecting the total number of delinquent behaviours engaged in. The parent report was identical and assessed young person delinquent behaviours to the extent that the parent was aware of them. Test-retest

¹ The official minimum age for FFT is 10 but a single 9 year-old was accepted by the FFT programme during the period of the study.

reliabilities for self-report delinquency measures are reported to be in the good to excellent range (0.84-0.97) for most groups and scoring procedures. Validity is reported to be well established in some contexts (ability to discriminate between groups, correlation between official and self-reported offenses) but not others (i.e. low correlation with arrests) (Huizinga & Elliott, 1986).

Family background

A selection of key indices of family functioning derived from the literature on the development of conduct problems are reported. Each is a percentage or count variable indicative of social and economic disadvantage.

The Global Therapist Rating scale

Therapist competence and model-adherence was assessed every three months using the Global Therapist Rating Scale (GTRS: Sexton & Alexander, 2004). This measure included the quarterly average of the programme supervisor's weekly ratings of therapist dissemination fidelity and clinical fidelity as well as a checklist of 38 "general clinical skills". Dissemination fidelity related to tangible tasks such as attendance at meetings and completion of paperwork. Clinical fidelity consisted of "clinical adherence" to the FFT model and "clinical competence" in the delivery of this model, where a minimum level of adherence was regarded as a necessary condition for competence. Dissemination fidelity was rated on a seven point scale from "none" to "always". Clinical fidelity was the sum of two four point scales; "clinical adherence" rated from "none" to "extensive/consistent" and "clinical competence" rated from "none" to "high". The 38 skills were each specific to one of five phases or aspects of FFT treatment; engagement, behaviour change, generalisation, general FFT skills and relationship skills. For instance an engagement and motivation skill was *Does the therapist refrain from taking sides, blaming or judging family members?* A behaviour change skill was *Does behaviour change begin only after successful progress towards engagement and motivation goals?* Each skill was rated on a five point scale from "never" to "always". The two fidelity scores and the average of the 36 competence scores were taken as representing the therapists'

fidelity and competence for those cases treated that quarter. Where a particular case spanned two quarters, the means of the three ratings were used.

Parent Satisfaction Survey

An eleven item survey was created based on the Client Satisfaction Questionnaire (CSQ-8: Larsen, Attkisson, Hargreaves & Nguyen, 1979) with item content modified to reflect considerations specific to FFT. An additional three items were created for Māori parents to assess the cultural acceptability of the FFT treatment and FFT therapists.

Results

An intention to treat approach was taken and all participants who commenced treatment were assessed whether or not they completed treatment, with treatment commencement defined as at least one FFT treatment session. There was a significant loss of post and follow-up data due largely to families dropping out of treatment and becoming uncontactable or refusing further assessments; 59 pre, 44 post and 47 follow-up parent interviews were obtained. Numbers for the teacher-report measure were limited by the high proportion of young persons not attending school. Drop-out was defined as a) a unilateral decision by the parents to discontinue treatment, as reported by the therapist and b), completion of two or less sessions. Using this criterion 8 or 13% of participants dropped out of treatment. The mean number of FFT sessions attended was 10.4 with a range of 1 to 31.

Conduct Problems

Multiple one-way repeated measures analysis of variance were employed. A test for linear trend representing the effect of time, or progressive improvement, was conducted with planned comparisons of pairwise differences between means. Alpha was set at $p=0.05$ overall or $p = 0.01$ per measure on the basis of a Bonferroni adjustment for five tests. Owing to a small number of extreme values the delinquent activities and child alcohol and drug measures were truncated at a maximum value of 100. In order to facilitate comparison all measures were standardised with a mean of 100 and a standard deviation of 10. Effect sizes were calculated as Cohen's d

with pairwise deletion of missing values. Values of d are described as follows; $d \geq 0.20$ = small, $d \geq 0.50$ = medium and $d \geq 0.80$ = large. STATA versions 12 and 13 and G*Power 3.1.5 were used for all statistical calculations.

Table 1 presents effect sizes and tests of statistical significance for five anti-social behaviour outcomes. The table shows the following:

i. Parent reported Conduct Disorder and Oppositional Defiant Disorder behaviours (CD/ODD) reduced from pre-test to follow-up to a degree consistent with a medium to large effect size ($d = 0.78$) and a test for linear trend indicating progressive improvement across assessments was statistically significant ($p < 0.0001$). Pairwise comparisons indicated that mean scores at both post-test and follow-up were significantly lower than scores at pre-test ($p < 0.01$) indicating a significant reduction in CD and ODD behaviours at post-test and follow-up relative to pre-test.

ii. Teacher reported Conduct Disorder and Oppositional Defiant Disorder behaviours reduced to a degree consistent with a medium effect size at follow-up ($d = 0.71$) and a test for linear trend was not significant ($p > 0.08$). No pairwise comparisons were significant. Achieved power for the linear trend was calculated as 0.84, an adequate value, indicating that the nonsignificant results were not due to low statistical power.

iii. Parent reported delinquent behaviours improved from pre-test to follow-up to a degree consistent with a small to medium effect size ($d = 0.49$) and the test for linear trend was not significant ($p > 0.03$). The contrast between pre-test and follow-up was significant ($p < 0.01$) indicating a significant reduction in parent-reported delinquent behaviours at follow-up relative to pre-test. Achieved power for the linear trend was calculated as 0.81, an adequate value, indicating that the nonsignificant results were not due to low statistical power.

iv. Young person reported delinquent behaviour improved to a degree consistent with a medium effect size at follow-up ($d = 0.68$) and the test for trend was significant ($p < 0.0013$). Contrasts between pre-test and both post-test and follow-up were significant (both $p < 0.002$) indicating that the level

of self-reported delinquent behaviours was significantly lower at post-test and follow-up compared to pre-test.

v. Young person alcohol and drug use, as reported by the parent, did not change; the effect size was close to nil ($d=0.08$) and a test for linear trend was nonsignificant ($p=0.75$). No contrasts were significant ($p>0.45$). Achieved power for the linear trend was calculated as 0.09, very low, thus the test possessed insufficient statistical power to detect the obtained change. However the low power was likely to be due to the small effect size hence the conclusion remains the same; young persons' level of alcohol and substance use did not change or improve over the course of treatment and follow-up.

vi. To test the overall significance of the findings in Table 1 we fitted a Multivariate Analysis of Variance (MANOVA) model to the data. In this model missing data were estimated using a multiple imputation approach (Rubin, 1987). Teacher CD/ODD data was omitted as the available data was too incomplete to warrant imputation. The analysis found a general trend for outcome measures to improve over time across the three measurement points ($F[8,50]=8.48, p<.0001$, Wilk's $\lambda=0.42$, partial $\eta^2=0.58$, power to detect the effect 1.0). To ensure that the results were not unduly influenced by the imputation and truncation procedures the MANOVA was repeated with a) no imputed data, b) non-truncated data and both a) and b); all analyses were statistically significant ($p<0.01$).

Table 1
Young person conduct problems: Tests for trend and pairwise contrasts with Cohen's d effect size for the pre to follow-up interval

Measure	n	Mean (sd) pre	Mean (sd) post	Mean (sd) follow-up	(df) F linear trend	p	Pre - Follow-up d (95% CI)	Contrasts $p<0.01$
CD/ODD (parent)	41/46	104.19 (9.62)	98.45 (9.79)	97.36 (9.43)	(2, 80) 13.61	0.0001*	0.78 (0.36 to 1.2)	pre vs. post pre vs. fup
CD/ODD (teacher)	20/22	103.28 (9.91)	98.95 (11.05)	96.76 (8.51)	(2, 38) 2.75	0.08	0.71 (0.10 to 1.32)	Nil
SRD (parent)	41/46	102.56 (11.71)	99.91 (10.29)	97.52 (7.83)	(2, 80) 3.75	0.028	0.49 (0.07 to 0.90)	pre vs. fup
SRD (young person)	26/32	104.72 (12.03)	97.93 (7.96)	97.34 (8.04)	(2, 50) 7.62	0.0013*	0.68 (0.17 to 1.18)	pre vs. post pre vs. fup
Alcohol/ drug use	41/46	100.54 (10.28)	99.42 (9.81)	100.04 (10.04)	(2, 80) 0.29	0.75	0.08 (-0.33 to 0.48)	Nil

Note. N is for linear trend/Cohen's d
* $p<.01$

Outcomes for Māori vs. non-Māori

Multiple t-tests were used to test for differences in mean pre to follow-up

change score (i.e. pre-test score – follow-up score) for each measure. Greater improvements for either group would be evidenced by a) larger effect sizes and b) larger change scores. Table 2 presents effect sizes and tests for identity of change scores for Māori and non-Māori for the previous measures.

Table 2.
Effect size for Māori and Non-Māori pre to follow-up and t-tests for differences in change scores

Measure	n	d (Māori)	95% CI	n	d (Non-Māori)	95% CI	t	p
CD/ODD ¹ (parent)	21	0.68	(0.05 to 1.29)	25	0.90	(0.32 to 1.48)	0.02	0.98
CD/ODD (teacher)	10	0.45	(-0.44 to 1.33)	12	0.93	(0.08 to 1.77)	1.08	0.29
SRD (parent)	21	0.19	(-0.42 to 0.80)	25	0.71	(0.13 to 1.28)	1.56	0.13
SRD (young person)	13	1.05	(0.21 to 1.86)	19	0.46	(-0.19 to 1.10)	1.18	0.25
Alcohol/drug use	21	0.24	(-0.36 to 0.85)	25	-0.04	(-0.59 to 0.51)	1.11	0.27

This table shows the following:

- i. There were no significant differences between mean change scores for Māori vs. non-Māori for any measure ($t 0.02 - 1.56, p 0.98 - 0.13$).
- ii. For all Māori/non-Māori pairs of effect sizes, confidence intervals for

between Māori and non-Māori effect estimates.

iii. At follow-up non-Māori achieved larger effects than Māori in three of five measures and Māori larger effects than non-Māori in the remaining two measures. The margins between effect sizes were variable with large differences

between estimates favouring both Māori and non-Māori. When a weighted average was calculated (weighted by n per measure), mean effects were similar with a slight advantage to non-Māori ($d=0.48$ vs. 0.56). The magnitude of the difference was minimal (0.08) and equivalent to less than a small effect (less than 0.20).

iv. A 2 x 3, group x time, MANOVA was fitted to provide a test for differential outcomes over time across the all four conduct problem measures combined (teacher CD/ODD omitted, missing data imputed). The test for interaction was not significant indicating that there was no detectable difference between the changes in conduct problems over time for Māori vs. non-Māori; $F(8,50)=0.53, p=0.83$, Wilk's $\lambda=0.92$, partial $\eta^2=0.08$, power to detect the effect 0.48. Whilst power was low (0.80 is the conventional minimum adequate power), Wilk's λ and partial η^2 indicated that only 8% of the variance in the measures was explained by progressive differences between Māori and non-Māori conduct

each measure overlapped considerably indicating a high degree of similarity

problems over time; this is consistent with the t-tests and effect sizes in Table 2 and suggests that differences in outcomes between Māori and non-Māori were minimal.

Treatment fidelity

Three therapist scores were assigned to each young person; fidelity, dissemination fidelity and average competence score. Ten participants who had completed less than four sessions were excluded in order to limit the analysis to those who had received more than minimal treatment. Participants were separated into high and low fidelity and competence groups on the basis of a median-split; three 2 x 3 group by time MANOVA's were conducted using a multiple imputation approach for missing data and excluding the teacher CD/ODD measure as before. The results were as follows:

i. All three MANOVAS failed to reach significance ($p > 0.05$) indicating that young persons treated with better than average fidelity or competence did not experience better outcomes in terms of reduced conduct problems than those young persons treated with lower levels of fidelity or competence; (fidelity [$F(8,40) = 0.32, p = 0.95, \text{Wilk's } \lambda = 0.94, \text{partial } \eta^2 = 0.06$, power to detect the effect 0.31], dissemination fidelity [$F(8,40) = 0.55, p = 0.81, \text{Wilk's } \lambda = 0.90, \text{partial } \eta^2 = 0.10$, power to detect the effect 0.51] and competence [$F(8,40) = 0.79, p = 0.62, \text{Wilk's } \lambda = 0.86, \text{partial } \eta^2 = 0.14$, power to detect the effect 0.68]). Whilst achieved power was low (< 0.69), values of Wilk's λ and partial η^2 indicated that only between 6% and 14% of the variance in outcomes was explained by the interaction between time and group, which is consistent with there being little if any difference in outcomes between the high and low fidelity/competence groups.

ii. Examination of the mean ratings for the low fidelity/competence groups suggests that the therapists' performance was poorer but not necessarily poor. For instance, the average score for the low clinical fidelity group was equivalent to a rating of "regular" or "frequent" adherence and "low" to "moderate" competence. The average for the low dissemination fidelity group was equivalent to a rating of "many" to "most" of dissemination adherence

procedures being followed. The average for the low competence group was equivalent to a rating of "sometimes" to "frequently" achieving each skill. Hence the failure to find a significant effect due to fidelity or competence may be due to the "low" groups manifesting an adequate level of fidelity and competency and achieving adequate client outcomes.

Parent satisfaction

Table 3 presents the percentage of parents responding to items in a parent satisfaction survey.

i. Parental satisfaction with FFT was generally high. 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. 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. 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. The mean level of satisfaction reported by Māori parents was compared to that of non-Māori; whilst the latter reported a slightly higher level of satisfaction overall a statistical test of this difference was not significant ($p > 0.66$).

Table 4 presents the responses of Māori parents to three items specific to the cultural acceptability of FFT 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. A similar proportion of parents felt that FFT was an appropriate intervention for Māori whānau.

Discussion

This study sought to evaluate the effectiveness and acceptability of FFT for Māori and non-Māori in New Zealand.

Table 3
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%

Note. Some rows may not sum to 100 due to rounding error

Table 4
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%

Note. Sample size = 20

Five hypotheses were tested relating to reductions in conduct problems, outcomes for Māori, parent satisfaction/cultural acceptability for Māori and the influence of therapist competence and fidelity. The first two hypotheses were supported to a moderate degree and the second two strongly supported. The last hypothesis was not supported. The hypotheses are discussed in turn.

In terms of the first hypothesis, two of five conduct problem measures showed statistically significant reductions at follow-up and medium effect sizes ($p < 0.01$, $d = 0.68, 0.78$). Two measures showed medium effect sizes but were not significant ($p > 0.028$, $d = 0.49, 0.71$). A fifth measure, alcohol and drug use, was not significant and showed minimal change equivalent to a small effect ($p > 0.75$, $d = 0.08$). A multivariate analysis which combined all five measures was statistically significant ($p < 0.001$).

As the predominant outcome measure used in the FFT literature has been officially recorded offending there are few studies using behaviour rating measures as in the present study. Graham et al. (2014) in an effectiveness and transportability study conducted in Ireland obtained pre to post effects of $d = 0.64$ for treatment completers ($n = 98$) using the conduct problems scale of the SDQ measure. Hartnett et al. (2015) in an Irish study with elements of efficacy, effectiveness and transportability (community based, however an FFT developer was co-author), obtained pre to follow-up effects of $d = 1.07$ and $d = 0.48$ for parent and youth rated conduct problems respectively. White et al. (2013) in a USA based community study obtained a pre-post effect of $d = 0.41$ for the conduct problems scale of the BASC rating scale ($n = 77$).

Three points can be made here: Firstly, Graham et al. (2014) and Hartnett et al. (2015) both conducted multivariate analyses of outcomes at post-test or follow-up and as in the present study these were statistically significant, indicating an overall reduction in conduct problems across measures. Secondly, the effect size of 0.08 for alcohol and drug use is an outlier in terms of one common definition; it is more than 1.5 times the interquartile range below the first quartile (cf. Howell, 1992, p51). This suggests that it may be a qualitatively different

outcome and not strictly comparable to the remaining four effect sizes. Were this the case the mean effect size for the current study and the three comparison studies is very similar; 0.66 range 0.49-0.78 (this study), mean 0.65 range 0.41-1.07 (comparison studies). Thirdly, effect sizes for the two nonsignificant outcomes (excluding alcohol and drug use) of 0.49 and 0.71 were within the range of effect sizes achieved in the three comparison studies, indicating, statistical significance aside, broadly comparable outcomes. This point holds despite the tests not being under-powered. These outcomes show that, notwithstanding the mixed results in terms of statistical significance and effect size, the reductions in conduct problems achieved are commensurate with outcomes obtained in previous FFT effectiveness and transportability studies.

With regards to the outcome for alcohol and drug use, previous studies of FFT have obtained significant results and medium to large effect sizes for reductions in alcohol and substance use (Rohde et al. 2014; Sleznick & Prestopnik, 2009). However substance use was a presenting issue in these studies and the youth exhibited high levels of this behaviour at the commencement of treatment. In contrast, substance use as the primary presenting issue was an exclusion criteria in the present study and initial levels of alcohol and substance use were low; 53% of young persons in the present study were rated by their parents as engaging in nil alcohol or substance use at the commencement of the study and a total of 84% of young persons were reported to engage in substance or alcohol use once a week or less. Thus there may have been a floor effect insofar as nil and low rates of use are difficult to improve upon. Lower rates of use may also have been related to the mean age of the young persons, 13 years 7 months whereas youth in the previously cited two studies were 15-16 years old on average and longitudinal studies in New Zealand show that rates of substance use increase across the teenage years (Fergusson et al., 2003). Thus the poor outcome on this measure may be unrelated to the effectiveness of the intervention.

The second hypothesis was that conduct problem outcomes for Māori would be similar in magnitude to those for non-Māori. In the present study

Māori achieved larger effect sizes than non-Māori for two of five measures and non-Māori larger effects than Māori on three measures. Statistical tests for differences in change scores and ethnicity by time interactions were not significant. Conclusions are limited by the small sample size and consequent low power, however on the basis of the present data outcomes for the two ethnic groups were not statistically discriminable. Further, the average effect sizes of 0.48 and 0.56 were clearly similar, the effect for Māori being 86% of that for non-Māori and the difference between effects (0.08) would be classified as a small effect. Overall the second hypothesis is supported. The international literature on this issue is mixed; in terms of the FFT literature comparisons have been confined to studies of alcohol & substance using youth and have found either no effect due to ethnicity or an advantage to minority groups (Flicker et al. 2008; Sleznick & Prestopnik, 2009). However van der Stouwe, Asscher, Stams, Dekovic and van der Laan (2014) in a meta-analysis of Multi-systemic Therapy (MST) found MST to be more effective with non-ethnic minority youth. In a New Zealand based study Sturrock & Grey (2013), in a study of the Incredible Years model found significant differences between Māori and non-Māori with the latter exceeding the former by effect sizes of between 0.14 and 0.29 for Conduct Disorder and Oppositional Defiant Disorder behaviours. These findings affirm that Māori may be susceptible to poorer outcomes in interventions for conduct problems but there is encouraging evidence from this study that effect-size differences in FFT may be minimal.

The third hypothesis was strongly supported; as indicated above Parent satisfaction was high with 70%-80% of parents indicating that they were "very much" satisfied with key aspects of the FFT intervention and therapists. Satisfaction with outcomes was less pronounced with 16% of parents reporting no satisfaction with the impact of FFT on their child's behaviour and ability to get on with others. This may reflect the imperfect nature of interventions for conduct problems and that a proportion of young persons will not have experienced any significant positive change. For instance, disregarding the alcohol and

drug measure for which outcomes were clearly poor, on the remaining four conduct problem measures between 17% and 39% of young persons either did not progress or deteriorated to some degree. Thus satisfaction with outcomes may accurately reflect variability in the outcomes obtained.

Hypothesis four was strongly supported with 80% to 90% of Māori parents reporting being “very much” satisfied with the cultural sensitivity of the FFT intervention. The latter is particularly important given the over-representation of Māori amongst at risk youth receiving FFT in New Zealand. Low satisfaction would suggest a degree of cultural insensitivity on the part of the therapist which is *prima facie* likely to damage families’ engagement with the treatment which would in turn affect outcomes.

Hypothesis five was not supported and there was no detectable influence of therapist fidelity or competence on outcomes. However the statistical analysis was lacking in power, and inspection of the scores suggests that fidelity and competence were largely adequate. Given that Graham et al. (2014) found that high-adherence therapists had significantly more favourable outcomes using a different measure of fidelity, the TAM (Therapist Adherence Measure), it is possible that the measure used in the present study failed to accurately discriminate low and high adherence. Or it may simply have been the case that levels of fidelity and competence were more than adequate and too uniform to lead to significantly different outcomes for participants treated with (relatively) high or low fidelity or competence. Future studies would benefit from comparing both measures.

An additional finding is the level of adversity experienced by the families involved in this study as detailed in a previous section. This shows the interconnectedness of socio-economic factors and conduct problems and suggests that conduct problem interventions should ideally be complemented with programmes to address such aspects of socio-economic disadvantage as may be amenable to change e.g. family violence, benefit dependence and low educational achievement.

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 were due primarily to the FFT intervention. Pre-post non-experimental designs have weak internal validity and any apparent treatment effect may be due to maturation, regression to the mean, mortality bias, non-specific treatment effects or a combination of these factors. A randomised controlled trial is required to control for these influences.

Secondly, it was not possible to collect official offending data. Whilst official data is not without its weaknesses it does connect the intervention with a quantifiable socially and politically important outcome. The use of offending data would permit a more thorough comparison with the existing FFT 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. 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 (60-90 minutes). Families were referred to the FFT programme by Child Youth & Family (social services) 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 unilaterally terminated treatment according to the FFT therapist. The most common manner of termination was family avoidance or withdrawal from treatment. Families unmotivated to complete further treatment sessions were almost always unmotivated to complete further interviews. The attrition rate amongst those families who completed treatment was 12% (5 of 40). Significantly, the attrition rate was very similar for Māori and non-Māori; a higher attrition rate for Māori could signal a problem with the

cultural acceptability of the intervention.

By way of comparison, treatment drop-out rates in the FFT literature range from 14% (Waldron et al., 2001) to 23% (White et al., 2013). Data loss rates may exceed drop-out rates where some participants who complete treatment nonetheless miss assessments; Hartnett et al. (2015) report three drop-outs in their treatment group but were able to obtain follow-up assessments for 53% of those pre-tested, a data-loss rate of 47%. Thus the rate of drop-outs and data-loss in the present study was not dissimilar to that seen in previous FFT studies.

Treatment drop-out is a ubiquitous issue in child psychotherapy and in particular with interventions for conduct problems and delinquency (Kazdin, 1990). In an earlier study of MST 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). A comparison of pre-test scores between participants with complete and incomplete data found somewhat larger (more severe) scores at pre-test in four of five conduct problem measures for those with incomplete data suggesting that more symptomatic participants were more likely to drop-out or miss assessments. However it was also found that those with higher scores at pre-test tended to obtain larger effect sizes, thus it is unclear whether any mortality bias served to attenuate or increase effect estimates.

In summary, the present study shows that FFT is effective and acceptable to both Māori and non-Māori families as implemented in a community setting in New Zealand. Effect sizes for conduct problem outcomes were favourable overall and comparable to those achieved in other FFT effectiveness and transportability studies. There were aspects of the design and implementation which temper these conclusions including the lack of a control group and a significant level of participant drop-outs. Ideally future

research should include a randomised controlled trial and collection of official offending data for at least two years post intervention. Loss of participants should be specifically addressed and modest compensation of parents for completing assessment interviews may be beneficial. If a randomised controlled trial is not possible additional single-group studies based in different centres in New Zealand would add greatly to the generality of the current findings.

References

- Alexander, J. F., & Parsons, B. V. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81*(3), 219-225.
- Baldwin, A., Christian, S., Berkeljon, A., Shadish, W & Bean, R., (2012). The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital and Family Therapy, Vol.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.
- Blisset, W., Church, J., Fergusson, D., Lambie, I., Langely, J., Liberty, K., et al. (2009). *Conduct problems: Best practice report*. Wellington: Ministry of Social Development.
- Cardemil, E. V. (2010). Cultural adaptations to empirically supported treatments: A research agenda. *The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology, Psychiatry, and Social Work, 7*(2), 8-21
- Carr, A. (2014). The evidence base for family therapy and systemic interventions for child-focused problems. *Journal of family therapy, 36*, 107-157.
- Cunningham, A. One step forward: Lessons learned from a randomized study of Multisystemic Therapy in Canada. *Praxis: Research from the centre for children & families in the justice system*. Retrieved from <http://www.lfcc.on.ca/onestep.html>
- 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.
- Curtis, N. M., Ronan, K. R., Heiblum, N., Reid, M & Harris, J. (2002). Antisocial behaviours in New Zealand Youth: Prevalence, interventions and promising new directions. *New Zealand Journal of Psychology, 31* (2), 53-58.
- Deater-Deckard, K., Dodge, K.A., Bates, J.E. & Pettit (1998). Multiple risk factors in the development of externalising behaviour problems: Group and individual differences. *Development and Psychopathology, 10*, 469-493.
- Department of Corrections. (2007). *Over-representation of Māori in the criminal justice system: An exploratory report*. Wellington, New Zealand: Department of Corrections.
- Elliott, D. S., Alexander, J., Pugh, C., Parsons, B. and Sexton, T. (2000). *Blueprints for Violence Prevention: Book Three Functional Family Therapy*. Boulder: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado.
- Elliott, D., & Huizinga, D. (1989). Improving self-reported measures of delinquency. In M.W. Klein (Ed.), *Cross-national research in self-reported crime & delinquency* (pp. 155-186). Boston, MA: Kluwer.
- Fergusson, D., Poulton, R., Horwood, J., Milne, B. & Swain-Campbell, N. (2003). *Comorbidity and coincidence in the Christchurch and Dunedin longitudinal studies*. Report prepared for the New Zealand Ministry of Social Development and Ministry of Education and the Treasury. Retrieved from <http://www.otago.ac.nz/christchurch/otago014857.pdf>
- Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, Vol.46*(8), Aug 2005, pp. 837-849.
- Fergusson, D. M., Horwood, L. J. & Stanley, L. (2013). Preliminary evaluation of the incredible years teacher programme. *New Zealand Journal of Psychology, Vol.42*(2), 2013, pp. 51-56.
- Fergusson, D.M., Stanley, L. & Horwood, L. J. (2009). Preliminary data on the efficacy of the incredible years basic parent programme in New Zealand. *Australian and New Zealand Journal of Psychiatry, Vol.43*(1), pp. 76-79.
- Flicker, S.M., Waldron, H.B., Turner, C.W., Brody, J.L. & Hops, H. (2008). Ethnic matching and treatment outcome with Hispanic and Anglo substance-abusing adolescents in family therapy. *Journal of Family Psychology, 22* (3), 439-447.
- Friedman, A.S. (1989). Family therapy vs. parent groups: Effects on adolescent abusers. *American Journal of Family Therapy, 17*, 335-347.
- 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 behaviour. *Criminal Justice & Behavior, 22*, 60-73.
- 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.
- Hallfors, D., Cho, H., Sanchez, V., Khatapoush, S., Kim, H. M., & Bauer, D. (2006). Efficacy vs effectiveness trial results of an indicated “model” substance abuse program: Implications for public health. *American Journal of Public Health, 96*(12), 2254-2259.
- Hansson, K. (1998). *Functional family therapy replication in Sweden: Treatment outcome with juvenile delinquents*. Paper presented to the Eighth International Conference on treating addictive behaviours. Santa Fe, New Mexico.
- Hansson, K., Johansson, P., Drott-Englen, G. & Benderix, Y. (2004). Funktionell familjeterapi i barnpsykiatrisk praxis. *Nordisk Psykologi, 56*(4), 304-320.
- Hartnett, D., Carr, D., & Sexton, T. (2015). The effectiveness of functional family therapy in reducing adolescent mental health risk and family adjustment difficulties in an Irish context. *Family Process, 10*, 1-18.
- 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.
- Henggeler, S. W., & Sheidow, A. J. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *Journal of Marital and Family Therapy, 38*(1), 30-58.
- Howell, D.C. (1992). *Statistical methods for psychology, 3rd ed*. Belmont, California:

- Duxbury Press.
- Huizinga, D. & Elliott, D. (1986). Reassessing the reliability and validity of self-report delinquency measures. *Journal of quantitative criminology*, 2, 293-327.
- Kazdin, A. E. (1990). Premature termination from treatment among children referred for antisocial behavior. *Child Psychology & Psychiatry & Allied Disciplines*, 31(3), 415-425.
- Kazdin, A.E. (2002). Psychosocial treatments for conduct disorder in children and adolescents. In P.E. Nathan & J.M. Gorman (Eds.), *A guide to treatments that work* (2nd ed., pp. 57-85). New York: Oxford University Press.
- Klein, N., Alexander, J., & Parsons, B. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting & Clinical Psychology*, 45, 469-474.
- Kratzer, L., & Hodgins, S. (1997). Adult outcomes of child conduct problems: A cohort study. *Journal of Abnormal Child Psychology*, 25(1), 65-81.
- Lantz, B.L. (1982). *Preventing adolescent placement through Functional Family Therapy and tracking*. Utah Department of Social Services, West Valley Social Services, District 2K, Kearns, UT 84118. Grant # CDP 1070 UT 83-0128020 87-6000-545-W
- Larsen, D.L., Attkisson, C.C., Hargreaves, W.A. and Nguyen, T.D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*, 2, 197-207.
- La Roche, M.J., & Christopher, M.S. (2008). Culture and empirically supported treatments: on the road to a collision? *Culture & Psychology*, 14, 333-356
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48(12), 1181-1209.
- Littell, J., Campbell, M., Green, S. & Toews, B. (2005). *Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17* Cochrane library, issue 3, Chichester, UK: John Wiley & Sons, Ltd.
- Marie, D. (2010). Maori and criminal offending: A critical appraisal. *Australian and New Zealand Journal of Criminology*, 43(2), 282-300.
- 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(9), 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.
- Marie, D., Fergusson, D. M., & Boden, J. M. (2014). Childhood socio-economic status and ethnic disparities in psychosocial outcomes in New Zealand. *Australian and New Zealand Journal of Psychiatry*, 48(7), 672-680.
- Menting, A. T., de Castro, B. O., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clinical Psychology Review*, 33(8), 901-913.
- Offord, D. R., & Bennett, K. J. (1994). Conduct disorder: Long-term outcomes and intervention effectiveness. *Journal of the American Academy of Child & Adolescent Psychiatry*, 33(8), 1069-1078.
- Owens, V. (2001). Whanake Rangatahi: Programmes and services to address Māori youth offending. *Social Policy Journal of New Zealand*, 16, 175-190.
- Robbins, M. (2016, April). Summary of research outcomes by study type: Efficacy, effectiveness, dissemination/implementation, basic and process [Presentation notes] *Dissemination for sustainability*. Paper presented at the Blueprints for violence prevention conference, Denver USA.
- 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*. Vol.82(2), pp. 342-348.
- Ryder, A. G., Ban, L. M., & Chentsova-Dutton, Y. E. (2011). Towards a cultural-clinical psychology. *Social and Personality Psychology Compass*, 5(12), 960-975.
- Rubin, D. B. (1987) *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Ryder, A. G., Yang, J., Zhu, X., Yao, S., Yi, J., Heine, S. J., et al. (2008). The cultural shaping of depression: Somatic symptoms in China, psychological symptoms in North America? *Journal of Abnormal Psychology*, 117(2), 300-313.
- Sexton, T., & Alexander, J. (2004). *Functional family therapy clinical supervision training manual*. Baltimore, MD : Annie E. Casey foundation.
- 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.
- Silva, M.D., Alpert, M., Pouget, E., Silva, V., Trosper, S., Reyes, K., & Dummit, S. (2005). A rating scale for disruptive behaviour disorders, based on the DSM-IV item pool. *Psychiatric Quarterly*, 76 (4), 327-339.
- 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. (2009) Comparison of family therapy outcome with alcohol-abusing, runaway adolescents. *Journal of Marital and Family Therapy*. Vol.35(3), pp. 255-277.
- Statistics New Zealand (2016). *Prison facts and statistics - march 2016*. Retrieved from http://www.corrections.govt.nz/resources/research_and_statistics/quarterly_prison_statistics/PS_March_2016.html#ethnicity
- Streiner, D. L. (2002). The 2 "Es" of Research: Efficacy and Effectiveness Trials. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie*, 47(6), 552-556.
- Sturrock, F., & Gray, D. (2013). *Incredible Years Pilot Study: Evaluation Report*. Wellington, NZ: Ministry of Social Development.
- Sundell, K., Hansson, K., Lofholm, C. A., Olsson, T., Gustle, L.-H., & Kadesjo, C. (2008). The transportability of multisystemic therapy to Sweden: Short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, 22(4), 550-560.
- Youth Horizons (2015) *Youth Horizons Annual report 2015*. Retrieved from <http://www.youthhorizons.org.nz/assets/Annual-Report-2015.pdf>
- van der Stouwe, T., Asscher, J. J., Stams, G. J. J., Dekovic, M., & van der Laan, P. H. (2014). The effectiveness of multisystemic therapy (MST): A meta-analysis. *Clinical Psychology Review*, 34(6), 468-481.
- 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.
- Waldron, H. B., & Turner, C. W. (2008). Evidence-based psychosocial treatments for adolescent substance abuse. *Journal of Clinical Child and Adolescent Psychology*, 37(1), 238-261.
- Washington state institute for public policy (2016). *Benefit cost results*. Retrieved from http://www.wsipp.wa.gov/BenefitCost/Pdf/1/WSIPP_BenefitCost_Juvenile-Justice

Acknowledgements:

We would like to thank the Tindall Foundation and Child Youth & Family for their generous support of this project.

Corresponding Author:

Charles Heywood
Youth Horizons Trust
42 Vestey Drive,
Mt Wellington
Auckland, New Zealand
Email: charlesh@youthorizons.org.nz