Graham, C., Carr, A., Rooney, B., Sexton, T. & Wilson Satterfield, L. (2014). Evaluation of functional family therapy in an Irish context. *Journal of Family Therapy*, 36, 20–38. doi: 10.1111/1467-6427.12028

Evaluation of Functional Family Therapy in an Irish Context.

Clare Graham[a], Alan Carr[b] Brendan Rooney[c], Tom Sexton[d], & Laura Rachel Wilson Satterfield[e]

a. Researcher, School of Psychology, University College Dublin, Ireland and Family Therapist at Families First, Clondalkin, Dublin, Ireland.

b. Professor of clinical psychology, School of Psychology, University College Dublin, and Family Therapist at the Clanwilliam Institute, Dublin Ireland.

c. Lecturer, Department of Technology and Psychology, Dún Laoghaire Institute of Art Design and Technology, Dublin

d. Professor of psychology at the Department of Counselling and Educational Psychology, and Director of the Centre for Adolescent and Family Studies, University of Indiana, Bloomington, Indiana, USA

e. Researcher at the Department of Counselling and Educational Psychology, and the Centre for Adolescent and Family Studies, University of Indiana, Bloomington, Indiana, USA

Submitted to JFT on 30 July 2012. Revised and resubmitted in 16 December 2012. Accepted for publication on January 18TH 2013.

Correspondence address: Alan Carr, Professor of Clinical Psychology, School of Psychology, Newman Building, University College Dublin, Belfield, Dublin 4, Ireland. E. alan.carr@ucd.ie. P. +353-1-716-8740

Email list: Thomas Lee Sexton <thsexton@indiana.edu>, Clare Graham <clare.graham@ucd.ie>, Laura Wilson Satterfield <wilsonlr@indiana.edu>, brendan.rooney@iadt.ie

Running head: Functional Family Therapy

ABSTRACT

In an Irish context we conducted a retrospective archival study of Functional Family Therapy for adolescents with behavioural problems. Strengths and Difficulties Questionnaire data were collected from 118 families at the beginning and end of therapy (at baseline and follow-up for dropouts) in a community-based clinic in a socially disadvantaged Dublin suburb. Analyses of improvement in mean scores and clinical recovery rates showed that outcome was associated with treatment completion and therapist adherence. Therapy completers treated by high-adherent therapists had the most favourable outcome. In contrast, the worst outcome occurred for dropouts. The outcome of cases treated by low-adherent therapists fell between these two extremes. These results show that FFT may be effectively implemented in an Irish context, and that the effectiveness of treatment is associated with families remaining in treatment for an average of 17 sessions, and receiving treatment from therapists who implement FFT with a high degree of fidelity.

PRACTITIONER POINTS

- Use Functional Family Therapy to treat adolescents with behavioural problems.
- To maximize benefits to clients, adhere to the Functional Family Therapy treatment protocol when conducting therapy.

INTRODUCTION

In Ireland, the context for the study reported in this paper, two large community surveys have shown that up to 20% of adolescents have significant behavioural problems (Lynch et al., 2004; Martin et al., 2006), a figure consistent with results of epidemiological studies of youth mental health problems in other countries (Costello, 2004; Ford, 2008). Family intervention programmes have shown particular promise in ameliorating adolescent behavioural problems, and Functional Family Therapy (FFT) has consistently been identified in authoritative international reviews as one such evidence-based programme (Baldwin et al., 2012; Henggeler & Sheidow, 2012). Few evidence-based family intervention programmes have been established in Ireland to address adolescent behavioural problems. In this paper a preliminary evaluation of FFT in an Irish context is described.

FFT is an evidence-based treatment for adolescent behavioural problems, conduct disorder, substance misuse and delinquency (Alexander & Parsons, 1982; Sexton, 2011). The FFT model has three distinct phases: engagement, behaviour change and generalization. Therapist goals and interventions appropriate to each phase are described in the treatment manual (Sexton & Alexander, 2004). The manualization of FFT has facilitated its dissemination internationally. When FFT is disseminated to community-based sites, adherence to the model (or treatment fidelity) is achieved through a process of intensive training and supervision. The FFT supervision process has also been detailed in a manual (Sexton, Alexander, & Gilman, 2004). Through telephone supervision with an expert FFT supervisor, therapists' adherence to the FFT model in community-based sites is assessed regularly. Client progress in community-based sites is tracked from session to session. Data on treatment fidelity and client progress are routinely entered by supervisors and therapists into the Functional Family Therapy Quality Improvement System (FFT-Q-System), which yields regular reports on model fidelity and therapy process and outcome. The FFT-Q system is a secure, web-based quality improvement information system.

A series of evaluation studies has shown that FFT is effective in reducing criminal activity by up to 60%, reducing treatment dropout from 50% to 20%, and improving family functioning in areas such as communication and problem-solving (Baldwin et al., 2012; Henggeler & Sheidow, 2012; Sexton, 2011). Furthermore, there is evidence that treatment fidelity mediates outcome in FFT, with cases treated by therapists who adhere to the model having better outcomes than those treated by low-adherent therapists, especially in cases at high risk due to family disorganization or deviant peer group membership (Barnoski, 2002; Sexton & Turner, 2010)

In 2007 a team of therapists at Families First, a community-based agency in a disadvantaged Dublin suburb in Ireland, was trained in FFT by Tom Sexton. Families First is part of Archways, a national organisation working in collaboration with local and national agencies to promote and research evidence-based programmes for children and young people. The provision of FFT at Families First is an innovation in the Irish mental health care system. This paper describes an evaluation of cases treated by the Families First FFT team. The evaluation was guided by the following two hypotheses: (1) that cases who completed a course of FFT would show significant improvement from intake to discharge; and (2) that cases who completed treatment with therapists who had high adherence to the FFT model would show greater improvement than dropouts or therapy completers treated by low-adherent therapists. We also wished to explore factors that predicted outcome, such as problem severity, age, gender and family composition.

METHOD

Design

This was a retrospective archival study of families who began a course of FFT at Families First between 2007 and 2011. To test the first hypothesis and assess improvement over the course of FFT from Time 1 (intake) to Time 2 (discharge), archival data collected from both parents and adolescents at initial and final therapy sessions from 98 families who

completed treatment were analysed. Therapist adherence data, collected at regular supervision sessions, was used to classify these 'treatment completers' into 49 families treated by therapists who showed high adherence to the FFT model, and 49 families treated by therapists with low model adherence. Improvement patterns in these two groups of cases were compared with that of a group of 20 cases who dropped out of treatment after 1 to 3 sessions. For these dropouts, archival data from first sessions (Time 1) were available. Follow-up (Time 2) data were collected by the first author over the telephone, between 9 and 46 months (mean = 23 months) after Time 1 data, from parents of families that dropped out of treatment. To test the second hypothesis, Time 1 and 2 data from dropouts and those treated by high- and low-adherent therapists were analysed. The drop-out group included all drop-outs for whom Time 1 data were available and who agreed to provide Time 2 data by telephone.

Participants

Demographic and referral characteristics of 20 dropouts, 49 cases treated by high-adherent therapists and 49 cases treated by low-adherent therapists are given in Table 1. There were no significant differences between groups on any of the variables listed in the table. Families were mainly of low socio-economic status with parents having semiskilled or unskilled occupations, or being unemployed (O'Hare, Whelan & Commins, 1991).

Table 1 here

Therapists and therapy

FFT was guided by a treatment manual (Sexton & Alexander, 2004) and conducted in families' homes or the Families First community-based treatment centre. The mean number of FFT sessions attended by families was 17 and therapy spanned a 3 to 6 month period. The mean numbers of sessions in each FFT phase were: engagement: 7,

behaviour change: 7 and generalisation: 5. There were 9 therapists in the study, 6 female and 3 male. Four had predominantly low Therapist Adherence Measure (TAM, Sexton et al., 2004) profiles with average annual TAM ratings lower than 3 (on a 7 point scale), and 5 had predominantly high TAM profiles with average annual TAM ratings of 3 or greater. All had primary degrees or postgraduate qualifications in mental health professions such as psychology, social work, psychotherapy, counselling or applied behavioural analysis. Each therapist received systematic training and ongoing supervision in FFT from Tom Sexton and Astrid Van Dam following the protocol detailed in the FFT supervision and treatment manuals (Alexander et al., 2000; Sexton & Alexander, 2004; Sexton et al., 2004). Therapists varied in the time they spent working on the project, and this ranged from 12 - 52 months. Case-loads of therapists varied from 1 to 29 cases. Numbers of treatment completers seen by therapists ranged from 1-26 and numbers of dropouts ranged from 0 to 7. There was no statistically significant association between therapist adherence (defined as having a predominantly high or low average annual TAM rating) and the proportion of completers and dropouts on therapists' case loads or the numbers of FFT sessions completers received. Seventy-four per cent of high-adherent therapists' cases completed therapy, and 87% of low-adherent therapists' cases were therapy completers. The mean number of sessions per case was 16.9 (range = 5-45) for high-adherent therapists and 17.3 (range = 7-36) for low-adherent therapists. There was also no statistically significant association between therapist gender or clinical experience and adherence to the FFT treatment model. While therapist experience was not associated with outcome on the total difficulties scale of the parent and adolescent versions of the Strengths and Difficulties Questionnaire (SDQ, Goodman 2001), therapist gender affected outcome. Cases treated by female therapists had better outcomes than those treated by males. On the total difficulties scale of the parent version of the SDQ, an ANOVA yielded a

significant Gender X Time interaction ($F(1, 96) = 4.92$; $p<0.05$). The mean score of adolescents from families treated by female therapists decreased significantly from 19.2 (SD = 6.0) at Time 1 to 15.0 (SD = 6.3) at Time 2. In contrast, the mean scores of those treated by male therapists were 19.6 (SD = 4.4) and 18.9 (SD = 6.3) at Times 1 and 2 respectively and these did not differ significantly from each other.

Instruments

Therapist adherence to the FFT model was assessed with the Therapist Adherence Measure (TAM, Sexton, Alexander, & Gilman, 2004). Adolescent behaviour problems were evaluated with parent and adolescent versions of the Strengths and Difficulties Questionnaire (SDQ, Goodman, 2001).

Therapist Adherence Measure. The TAM is a supervisor-rated index of FFT treatment fidelity (Sexton et al., 2004). In a clinical supervision telephone meeting with a therapist about a specific case, general adherence and phase-specific adherence are rated using 7-point Likert scales ranging from 0 = low adherence, through 3 = average adherence, to 6 = high adherence. These are averaged to give an overall TAM rating. General adherence, rated on a single Likert scale, is the degree to which supervisors perceive therapists to be following the FFT clinical model in the specific case presented during that clinical supervision discussion. Phase-specific adherence is the degree to which supervisors perceive therapists to be focusing treatment on the goals of the specific phase of the FFT clinical model in which the therapy is currently occurring. For each of the three FFT phases (engagement, behaviour change, or generalization) 6 Likert scale ratings are made. For the engagement phase ratings are made for alliance building, reframing, validating, supporting, facilitating expressing responsibility, and focusing the session. For the behaviour change phase ratings are made for targeting behaviours for change, linking targets to presenting problems, linking targets to organizing themes, applying behavioural skills, matching behaviour change skills to family style, and focusing the session. For the generalization phase ratings are made for selecting generalization

targets, facilitating autonomous skills use within the family, facilitating autonomous skill use outside the family (in school and community), facilitating planning to manage future challenges, matching generalization skills to families' styles and organizing themes, and focusing the session. The TAM supervisor rating scale has been adapted from videotape adherence rating systems which have shown high inter-rater reliability (Sydnor, 2006; Gilman, 2008). Barnowski (2002) and Sexton and Turner (2010) found that TAM scores predicted lower recidivism in juvenile delinquents treated with FFT.

In the present study expert FFT supervisors made TAM ratings for the 9 therapists involved in the project. For each year of the project (2007 through 2011), each therapist was given an average TAM score for that year, based on the TAM ratings they received in all telephone supervision meetings that year. For each therapist this average TAM rating was then linked to each family the therapist treated that year. For each year of the project, therapists were classified as high-adherent if their average TAM scores were 3 or greater; otherwise they were classified as low-adherent.

Strengths and Difficulties Questionnaire. The SDQ is a 25-item behavioural screening instrument, for assessing children and adolescents (Goodman 2001). It yields scores for total difficulties, conduct problems, hyperactivity, emotional symptoms, peer problems and prosocial behaviour scales. Three point response formats are used for all items (0 = not true, 1 = somewhat true, 2 = certainly true). There are parent and adolescent versions of the SDQ, and both have good psychometric properties (Goodman, 2001). In the present study parent and adolescent treatment completers filled out appropriate versions of the SDQ at Times 1 and 2. For dropouts, only the parent version was completed at Times 1 and 2. Cronbach's alpha reliability coefficients for scales from the parent and adolescent versions of the SDQ at Times 1 and 2 ranged from .58 to .82. In almost all cases, mothers (rather than fathers) completed the parent version of the SDQ. In all instances (except for the adolescent version of the SDQ at Time 1 where alpha = 0.6) total difficulties scales yielded alphas greater than 0.7 indicating good internal consistency reliability on this scale. On other scales reliability was moderate to good (0.58 – 0.74).

Ethics

This study was conducted with ethical approval of involved institutions.

RESULTS

Mean improvement in treatment completers from Time 1 to Time 2

To test the first hypothesis, and assess the clinical significance of differences between Time 1 and 2 mean SDQ scores, a MANOVA followed by paired t-tests were conducted on all 12 scales from parent and adolescent versions of the SDQ. A one-way MANOVA revealed a significant multivariate main effect Wilks' $\lambda = .58$, $F(1, 97) = 5.19$, $p < .001$, partial eta squared = .42. Power to detect the effect was 0.99. Thus, hypothesis 1 was supported. Results of paired t-tests given in Table 2 showed that significant improvement from Time 1 to 2 occurred on all SDQ scales, except the peer problems scale of the adolescent version of the SDQ. To control for type 1 error in these 12 analyses the rough false discovery correction was made (Benjamini & Hochberg, 1995). Effect sizes comparing means at Times 1 and 2 were computed using Cohen's (1988) formula (d = (Time 1 Mean – Time 2 Mean) / Pooled SD). From Table 2 it may be seen that effect sizes ranged from d = 0.12 - 0.94. A large effect size ($d > 0.8$) was found for parent-rated hyperactivity. A small effect ($d < 0.2$) occurred for adolescent-rated peer problems. Effect sizes for the remaining parent-rated scales and all of the adolescent-rated scales were in the moderate range (d = 0.2 - 0.8). Effect sizes for all parent-rated scales were larger than those for adolescent rated scales.

Table 2 here

Clinical improvement rates of treatment completers from Time 1 to Time 2

To examine clinical improvement rates, an issue relevant to the first hypothesis, rates of clinical improvement based on scores on the total difficulties scale of the parent and adolescent versions of the SDQ were determined in two ways. First, we calculated the

percentage of cases that scored below the clinical cut-off point after treatment, expressed as a function of the number of cases that scored above the clinical cut-off point before treatment. For these analyses clinical cut-off points on the total difficulties scale of 17 for the parent version and 20 for the adolescent version were taken from the SDQ website (http://www.sdqinfo.com/). Second, we calculated the percentage of all 98 treatment completers with a reliable change index (RCI) greater than 1.96. The RCI is an index of clinical improvement from one time point to another, which takes account of the psychometric properties of the scale used to assess the reliability of improvement. The reliable change index was calculated by subtracting SDQ total difficulties scores obtained at Times 1 and 2 and dividing this by the standard error of difference (Jacobson & Truax, 1991). The following equation was used to obtain the standard error of difference $\sqrt{2}$ (Standard Deviation $\sqrt{(1\text{-test-retest reliability})}$)^2. For the total difficulties scale of the parent version of the SDQ the standard error of difference was 4.34 based on a standard deviation of 5.8 in the normative sample (Meltzer et al., 2000, http://www.sdqinfo.com/) and a test-retest reliability of 0.72 (Goodman, 2001). For the total difficulties scale of the adolescent version of the SDQ the standard error of difference was 4.55 based on a standard deviation of 5.2 in the normative sample (Meltzer et al., 2000, http://www.sdqinfo.com/) and a test-retest reliability of 0.62 (Goodman, 2001).

Sixty-three of 98 treatment completers had Time 1 scores at or above the clinical cut-off score of 17 on the total difficulties scale of the parent version of the SDQ. Of these 63, 25 scored below the clinical cut-off at Time 2, indicating an overall clinical improvement rate of 39.68% from intake to discharge using this SDQ cut-off criterion.

Twenty-four of 98 treatment completers had Time 1 scores at or above the clinical cut-off score of 20 on the total difficulties scale of the adolescent version of the SDQ. Of

these 24, 10 scored below the clinical cut-off at Time 2, indicating an overall clinical improvement rate of 41.66% from intake to discharge using this SDQ cut-off criterion.

Sixteen of 98 treatment completers obtained RCIs greater than 1.96 on total difficulties scale of the parent version of the SDQ, representing an improvement rate of 16.32% using this conservative RCI criterion.

For the adolescent version of the SDQ 11 of 98 treatment completers obtained RCIs greater than 1.96, representing an improvement rate of 11.22% using this conservative RCI criterion.

Mean improvement of dropouts and cases treated by high and low adherent therapists.

The second hypothesis was that cases who completed treatment with therapists who had high adherence to the FFT model would show greater improvement than dropouts or therapy completers treated by low-adherent therapists. To test this hypothesis, a 3 X 2, Groups X Time MANOVA followed by a series of 3 X 2, Groups X Time ANOVAs were conducted on all scales from the parent version of the SDQ, with the rough false discovery correction for type 1 error and a study-wise p value of .05. In these analyses there were three groups: 49 cases treated by high-adherent therapists with TAM scores of 3 or greater; 49 cases treated by low-adherent therapists with TAM scores less than 3; and 20 dropouts who attended 3 or fewer appointments. For these 3 groups SDQ data collected at intake (Time 1) and discharge from treatment for completers, or 9 – 46 months after intake for dropouts (Time 2) were analysed. In these analyses the significant Groups X Time interactions were of central interest, since they indicated that the pattern of improvement or deterioration from Time 1 to 2 differed across the 3 groups.

The MANOVA yielded a significant Group X Time interaction, Wilks' $\lambda = .702$, F (2, 115) = 3.54, p <. 001, partial eta squared = .162. Power to detect the effect was .99. From Table 3 it may be seen that in a series of ANOVAs significant Group X Time interactions

occurred for all SDQ scales except the peer problems scale. These interactions are graphed in Figure 1. Tests of simple effects confirmed the impression given by Figure 1.

For cases treated by high-adherent therapists means at Times 1 and 2 on the total difficulties, conduct problems, hyperactivity, emotional problems and prosocial behaviour scales differed significantly, indicating that improvement on these scales occurred in this group. In contrast, for dropouts and cases treated by low-adherent therapists, means at Times 1 and 2 on these 5 SDQ scales did not differ significantly, indicating that no improvement occurred on any of these scales in these two groups.

Furthermore, at Time 2, means of the group treated by high-adherent therapists were significantly lower than those of the group treated by low-adherent therapists and dropouts on the total difficulties, conduct problems, hyperactivity, and emotional symptoms scales. These results indicate that the group treated by high-adherent therapists showed greater improvement after treatment than cases treated by low-adherent therapists and dropouts on these 4 scales. On the prosocial behaviour scale, at Time 2 the mean of the group treated by high-adherent therapists was significantly greater than that of dropouts, indicating that on this scale the group treated by high-adherent therapists showed greater improvement than dropouts at Time 2.

At Time 2 means of the group treated by low-adherent therapists were significantly lower than those of dropouts on the total difficulties, conduct problems, and hyperactivity SDQ scales. These differences largely reflect deterioration in the dropout group.

Effect sizes were computed for the 5 SDQ scales on which significant Groups X Time interactions were found in the ANOVAs reported above. Effect sizes at Time 2 for groups treated by high- and low-adherent therapists were computed by comparing means of these two groups at Time 2 with means of dropouts using Cohen's (1988) effect size formula (d = (Group 1 Mean – Group 2 Mean) / pooled SD). From Table 3 it may be seen that effect sizes for the group treated by high-adherent therapists were greater than those for the group treated by low-adherent therapists for the total difficulties, conduct problems, hyperactivity, emotional symptoms and prosocial behaviour scales. Effect sizes for the group treated by high-adherent therapists ranged from d = 0.65 - 1.59. In this group effect sizes for the total difficulties, conduct problems, hyperactivity and emotional problems

scales were in the large range (d > 0.8), and the effect size for prosocial behaviour was in the medium range (d = 0.2 - 0.8). In contrast, effect sizes for the group treated by low-adherent therapists ranged from d = 0.24 - 0.88. Only the effect size for the hyperactivity scale was in the large range (d > 0.8) and the remainder were in the medium range (d = 0.2 - 0.8). The results of the MANOVA, ANOVAs and effect size analyses support the second hypothesis.

Figure 1 here

Clinical improvement rates of dropouts and cases treated by high and low adherent therapists.

Clinical improvement rates of cases treated by high- and low-adherent therapists and dropouts, based on scores on the total difficulties scale of the parent version of the SDQ, were determined in the two ways described in a previous section. First we calculated the percentage of cases that scored below the clinical cut-off point of 17 on the parent version of the SDQ after treatment, expressed as a function of the number of cases that scored above the clinical cut-off point before treatment. Second we calculated the percentage of cases with an RCI greater than 1.96. Chi square tests were used to assess the statistical significance of differences in improvement rates in the three groups.

Using the SDQ clinical cut-off criterion, the improvement rate of the group treated by high-adherent therapists was 59.4% (19/32). This was significantly greater than that of the rates for the group treated by low-adherent therapists (19.4% (6/31)) and dropouts (0% (0/15)) (Chi square (2, N = 78) = 20.34, $p < .001$).

Using the very conservative RCI > 1.96 criterion, the improvement rate of the group treated by high-adherent therapists was 22.45%. This was not significantly greater than that of the rates for the group treated by low adherent therapists (10.20%) and dropouts (5.00%) (Chi square (2, N = 118) = 4.70, $p > .1$). These findings on improvement rates partially support the second hypothesis.

Exploratory regression analyses

Two exploratory step-wise multiple regression analyses were conducted to investigate the extent to which TAM scores, Time 1 scores from parent and adolescent versions of the SDQ, adolescent age, adolescent gender (female = 1, male = 2) and family composition (one parent family = 1, two parent family = 2) predicted Time 2 scores on the total difficulties scales of the parent and adolescent version of the SDQ for 98 treatment completers. Two predictors explained 39.1% of the variance in Time 2 parent SDQ total difficulties scores (R^2= .39, F (2, 97) = 30.53, p < .001). These were: Time 1 total difficulties scores from the parent version of the SDQ (β = .55, p < .001) and TAM scores (β = -.28, p < .001). From parents' perspectives, better outcome occurred in cases with less severe problems at intake and treated by more adherent therapists. Four predictors explained 47.7% of the variance in Time 2 adolescent SDQ total difficulties scores (R^2 = .47, F (4, 97) = 21.23, p < .01). These were: Time 1 total difficulties scores from the adolescent version of the SDQ (β = .55, p < .001), Time 1 total difficulties scores from the parent version of the SDQ (β = .32, p < .01), adolescent age (β = -.24, p < .01), and adolescent gender (β = .24, p < .01). From adolescents' perspectives, better outcome occurred in younger girls with less severe problems at intake.

Table 3 here

DISCUSSION

The first hypothesis - that cases who completed a course of FFT would show significant improvement from intake to discharge - was supported. On all but one SDQ scale, significant improvement in group mean scores occurred from intake to discharge. Clinical recovery rates using the SDQ clinical cut-off criterion were approximately 40% using the SDQ clinical cut-off criterion. However, they were less than half this using the very conservative RCI criterion. The second hypothesis - that cases who completed treatment with therapists who had high adherence to the FFT model would show greater improvement than dropouts or therapy completers treated by low-adherent therapists -

was also supported. On all but one SDQ scale, significant improvements in mean scores of cases treated by high-adherent therapists occurred, whereas no such improvement occurred in mean scores of dropouts or cases treated by low-adherent therapists. Clinical recovery rates using the SDQ clinical cut-off criterion were almost 60% for therapy completers treated by high-adherent therapists, about a third of this (19.4%) for cases treated by low-adherent therapists, and 0% for dropouts. Using the very conservative RCI criterion, a similar pattern occurred, although improvement rates were lower and differences were not statistically significant. Analyses using the RCI and SDQ clinical cut-off criteria were used to provide more and less conservative estimates of clinical improvement. Unfortunately, comparative data from other family therapy outcome studies using these procedures are not available.

Exploratory regression analyses showed that, from parents' and adolescents' perspectives, better adjustment at intake predicted better outcomes. From parents' perspectives, greater therapist adherence was also associated with better outcome. From adolescents' perspectives, better outcomes occurred for younger girls.

These results show that FFT may be effectively implemented in an Irish context, and that the effectiveness of treatment is associated with families remaining in treatment for an average of 17 sessions, and receiving treatment from therapists who implement FFT with a high degree of fidelity. These findings are consistent with those of Barnowski (2002) and Sexton and Turner (2010) who found that both therapist-adherence and psychosocial risk factors are both associated with outcome. The number of FFT sessions in the present study is similar to that in other studies, where the range is from 12 sessions in mild cases to 30 sessions in more severe or complex cases (Sexton, 2011).

This study had all the limitations associated with a retrospective archival study reliant on self-report, child-focused data. Clearly the results of the study would be

strengthened if observational data or recidivism records had been used, and family functioning as well as adolescent behaviour were assessed. In testing the first hypothesis, a single group design was used with no control group. Thus the degree to which changes were due to maturation or other developmental factors could not be determined. In testing the second hypothesis, while a three group design was used, with cases in the 3 groups differing in the amount and quality of FFT received, cases were not randomly allocated to these conditions. Cases self-selected to complete treatment or to drop out. Characteristics of completers and dropouts (for example degree of family disorganization, degree of deviant peer group membership, and extent of adolescent or parental personal vulnerabilities) that led to self-selection may also have determined differences in the outcome of these groups. While cases did not self-select high- or low-adherent therapists, there may have been some systematic bias in the allocation of cases to therapists which also accounted for the differing outcomes of cases in these two groups. Also the number of months between Time 1 and Time 2 assessments for drop-outs was greater than that for treatment completers, which may have accounted for dropouts' poorer SDQ scores. On the positive side, the 3 groups formed to test the second hypothesis did not differ on baseline demographic, clinical or referral characteristics such as the adolescents' age, gender, SDQ total difficulties scores, family composition, reason for referral and source of referral. The similarity of the groups on these variables reduces the possibility that extraneous variables may have accounted for the differing outcome of the three groups. However, because family processes were not assessed, it was not possible to determine if some families allowed clinicians to be more adherent to the FFT model and other families made it more difficult. This is an issue deserving investigation in future research.

A prospective randomized controlled trial which included both self-report and observational measures and assessed changes in family functioning as well as adolescent

behaviour would overcome the limitations of the present study. Such a study is currently underway.

REFERENCES

Alexander, J. & Parsons, B. (1982). *Functional family therapy: principles and procedures.* Carmel, CA: Brooks & Cole.

Alexander, J., Pugh, C., Parsons, B. & Sexton, T. (2000). Functional family therapy (Second Edition). In D. Elliott (Series Editor), *Blueprints for violence prevention* (Book 3). Boulder, CO: Centre for the Study and Prevention of Violence, Institute of Behavioural Science, University of Colorado.

Baldwin, S., Christian, S., Berkeljon, A. & Shadish, W. (2012). The effects of family therapies for adolescent delinquency and substance abuse: a meta-analysis. *Journal of Marital and Family Therapy,* 38, 281-304.

Barnoski, R. (2002). *Washington State's implementation of functional family therapy for juvenile offenders: Preliminary findings.* Olympia, WA: Washington State Institute for Public Policy.

Benjamini, Y. & Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1), 289–300.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (Second edition). Hillsdale, NJ: Erlbaum.

Costello, E. J., Mustillo, S., Keeler, G., & Angold, A. (2004). Prevalence of psychiatric disorders in childhood and adolescence. In B. L. Levin, J. Petrila & K. D. Hennessy (Eds.), *Mental health services: A public health perspective* (Second Edition, pp. 111-128). New York, NY, US: Oxford University Press.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Ford, T. (2008). Practitioner Review: How can epidemiology help us plan and deliver effective child and adolescent mental health services? *Journal of Child Psychology and Psychiatry,* 49, 900-914

Gilman, L. (2008). *Supervisory interventions and treatment adherence: An observational study of supervisor interventions and their impact on therapist model adherence.* (Unpublished doctoral dissertation). Bloomington, IN: Indiana University,

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *American Journal of Child and Adolescent Psychiatry*, 40, 1337-1345.

Henggeler, S. & Sheidow, A. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *Journal of Marital and Family Therapy,* 38, 30-58.

Jacobson, N. & Truax, P. (1991) Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59, 12-19.

Lynch, F., Mills, C., Daly, I. & Fitzpatrick, C. (2006). Challenging times: Prevalence of psychiatric disorders and suicidal behaviours in Irish adolescents. *Journal of Adolescence*, 29, 555-73.

Martin, M., Carr, A. & Burke, L., Carroll, L. & Byrne, S. (2006). *The Clonmel Project. Mental health service needs of children and adolescents in the south east of Ireland: final report.* Clonmel: Health Service Executive.

Meltzer, H., Gatward, R., Goodman, R. & Ford, F. (2000). *Mental health of children and adolescents in Great Britain*. London: The Stationery Office.

O'Hare, A., Whelan, C. & Commins, P. (1991). The development of an Irish census-based social class scale. *The Economic and Social Review*, 22, 135-156.

Sexton, T. (2011). *Functional family therapy in clinical practice*. New York: Routledge.

Sexton, T. & Alexander, J. (2004). *Functional family therapy clinical training manual*. Baltimore, MD: Annie E. Casey Foundation.

Sexton, T., Alexander, J. & Gilman, L. (2004). *Functional family therapy clinical supervision training manual*. Baltimore, MD: Annie E. Casey Foundation.

Sexton, T. & Turner, C. (2010). The effectiveness of Functional Family Therapy for youth with behavioural problems in a community practice setting. *Journal of Family Psychology*, 24, 339–348.

Sydnor, A. (2006). Assessing therapist adherence from video recordings using the TAM. Bloomington: Indiana University.

TABLE 1 *Demographic and referral characteristics of dropouts and cases treated by high and low adherent therapists*

	High adherence (n = 49)		Low adherence (n = 49)		Dropouts (n = 20)		Chi-square or F
	f	%	f	%	f	%	
Gender							5.43
Male	34	69.4	23	46.9	13	65	
Female	15	30.6	26	53.1	7	35	
Age							3.07
Mean	14.2		13.9		15.15		
Standard deviation	2.03		1.83		1.75		
Family compositions							2.52
Two-parent	25	51	17	34.7	8	40	
Single-parent	24	49	31	63.3	12	60	
Other	0	0	1	2	0	0	
Reason for referral							4.78
Family relationship difficulties	21	42.9	20	40.8	10	50	
School difficulties	15	30.6	12	24.5	4	20	
Aggressive behaviour	3	6.1	8	16.3	5	25	
Parenting issues	3	6.1	4	8.2	0	0	
Substance use	2	4.1	2	4.1	0	0	
Self-harm	1	2.0	0	0	1	5	
Other	4	8	3	6.1	0	0	
Source of referrals							8.98
Schools	28	57.1	20	40.8	9	45	
Mental health services	13	26.5	15	30.6	4	20	
Community agencies	4	8.2	7	14.3	3	15	
Youth Justice	1	2	2	4.1	4	20	
County Council	3	6.1	2	4.1	0	0	
Other	0	0	3	6.1	0	0	
Parent SDQ total difficulties score at Time 1							
Mean	19.02		19.51		20.85		0.70
Standard deviation	5.30		5.78		5.86		

Note. None of the Chi-square of F values are statistically significant at $P < 0.029$, which is equivalent to a study-wise P value of 0.05 using the rough false discovery correction for type 1 error. f, frequency.

TABLE 2 *Status of treatment completers on the parent and adolescent versions of the strengths and difficulties questionnaire (SDQ) at Time 1 and Time 2*

		Parent version of SDQ				Adolescent version of SDQ			
Variable		Time 1	Time 2	t	d	Time 1	Time 2	t	d
SDQ total difficulties	M	19.26	15.65	6.21**	.59	16.90	14.58	4.24**	.41
	SD	5.78	6.39			5.11	6.19		
SDQ conduct problems	M	5.16	3.75	6.43**	.64	4.48	3.63	3.72**	.43
	SD	2.26	2.15			1.89	2.03		
SDQ hyperactivity	M	6.24	2.30	3.50**	.94	5.52	4.96	2.40*	.22
	SD	5.39	2.41			2.37	2.61		
SDQ emotional symptoms	M	5.06	3.94	4.39**	.46	4.32	3.61	3.24*	.27
	SD	2.29	2.58			2.71	2.53		
SDQ peer problems	M	3.06	2.55	2.46**	.22	2.57	2.33	1.11	.12
	SD	2.55	2.10			2.15	1.81		
SDQ prosocial behaviour	M	6.47	7.24	3.63**	.35	6.64	7.30	3.19**	.34
	SD	2.30	2.08			1.90	1.99		

Note. *$P < 0.027$, which is equivalent to a study-wise P value of 0.05 using the rough false discovery correction for type 1 error, **$P < 0.01$. $N = 98$. d, effect size; M, mean; SD, standard deviation; t, value from t-test; Time 1, Intake; Time 2, discharge.

TABLE 3 *Status of dropouts and cases treated by high-adherent and low-adherent therapists on the parent version of the strengths and difficulties questionnaire (SDQ) scales at Time 1 and Time 2*

		High adherence (n = 49)		Low adherence (n = 49)		Dropouts (n = 20)		ANOVA F values			Effect sizes (d) at Time 2	
Variable		Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Groups	Time	G × T	High-adherent vs dropouts	Low-adherent vs dropouts
SDQ total difficulties	M	19.02	13.46	19.51	17.83	20.85	22.65					
	SD	5.30	5.78	6.27	6.27	5.86	5.79	8.16**	10.98**	14.49**	1.59	0.79
SDQ conduct problems	M	5.18	3.34	5.14	4.16	4.90	5.80					
	SD	1.90	1.94	2.60	2.29	2.61	2.62	2.10	8.49**	11.18**	1.07	0.67
SDQ hyperactivity	M	6.04	4.81	6.44	5.98	7.10	8.15					
	SD	2.11	1.42	2.49	2.29	2.22	2.62	8.48**	0.86	6.95**	1.59	0.88
SDQ emotional symptoms	M	5.12	3.22	5.00	4.67	5.10	5.25					
	SD	2.61	2.37	1.94	2.60	2.67	2.17	2.17	7.40*	6.79**	0.89	0.24
SDQ peer problems	M	2.93	2.08	3.18	3.02	3.30	3.50					
	SD	2.23	2.04	2.48	2.06	2.61	1.50	1.95	1.58	2.15	—	—
SDQ prosocial behaviour	M	5.93	7.30	7.02	7.18	5.95	5.80					
	SD	2.14	2.02	2.34	2.15	2.28	2.58	2.85	4.78	5.63*	0.65	0.58

Note. *$P < 0.029$, which is equivalent to a study-wise P value of 0.05 using the rough false discovery correction for type 1 error. **$P < 0.01$. $N = 118$. F values are from 3 × 2 Groups × Time ANOVAs. G × T, groups × time; M, mean; SD, standard deviation; Time 1, Intake; Time 2, discharge or 9–46 months after intake in the case of dropouts.

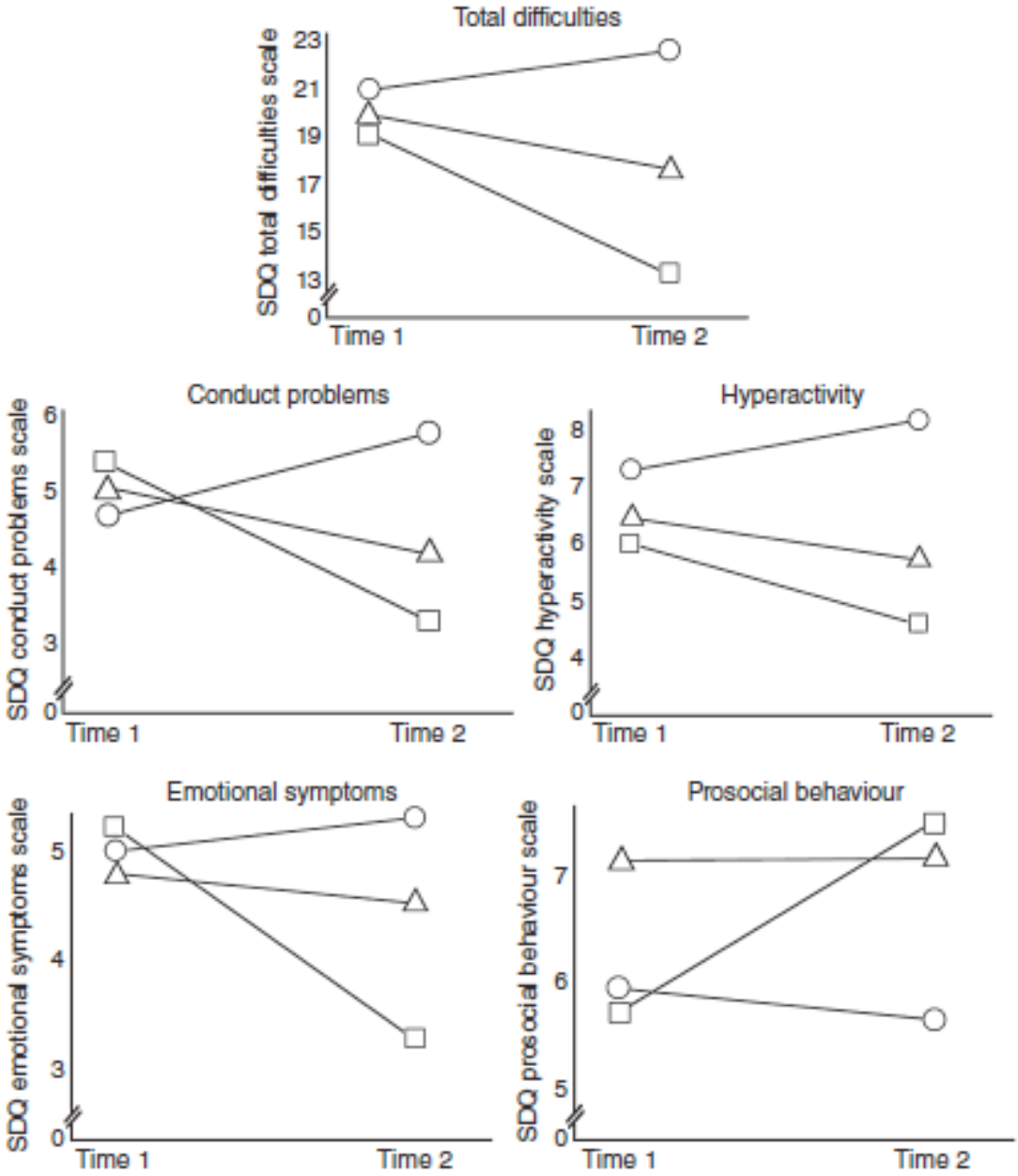

Figure 1. Status of ○ dropouts and cases treated by □ high-adherent and Δ low-adherent therapists on the parent version of the strengths and difficulties questionnaire (SDQ) scales at Time 1 and Time 2.