Complex interventions for children and young people: exploring service delivery frameworks and characterising interventions

Cressida Gaffney

D.Clin.Psy. thesis (Volume 1), [2012]

University College London

Overview

This thesis investigates the service delivery frameworks which support complex interventions for children and young people with conduct problems, and their families.

Part 1, the literature review, evaluates existing measures and other literature in the field to inform the development of a fidelity measure for the service delivery frameworks supporting complex interventions. 35 papers are examined using an approach informed by narrative synthesis to bring together the emerging themes. The service delivery frameworks which underpin interventions are little evaluated in the literature, and the review concluded that there is scope for the development of a measure to examine the service delivery elements of interventions for children and young people with conduct problems, which might be best informed by drawing on existing measures and literature on effective delivery of complex interventions.

Part 2, the empirical paper, describes the development and administration of the Children and Young People – Resource, Evaluation and Systems Schedule (CYPRESS) as part of the Systemic Therapy for At Risk Teens (START) randomised controlled trial (RCT) comparing multisystemic therapy (MST) with management as usual (MAU). CYPRESS was developed on the basis of a review of existing measures in the field, as well as research into the central aspects of service delivery which support complex interventions, as an interview-based measure of the service delivery frameworks supporting complex interventions. CYPRESS was piloted with two non-START trial teams, and subsequently administered to 16 teams (8 MST and 8 MAU) taking part in the START trial. The results of these interviews were used to compare the service delivery elements supporting MST and MAU, and

to characterise the MAU services in the trial. The importance of further development and testing of CYPRESS is noted.

Part 3, the critical appraisal, addresses methodological considerations arising from the research, discusses implications of the work, and reflects on the process of carrying out the research, and the context in which research of this nature occurs.

TABLE OF CONTENTS

Appendices

List of tables

Literature review

Empirical paper

List of figures

Acknowledgements

I would like to thank the following people who participated in the pilot interviews and the trial interviews: Kim Lawson, Sigrid Mannes, Garry Richardson, Nina Stevens (AMASS Islington); Geoffrey Baruch, Moira Lamond, Amanda Singh (The Brandon Centre); and the START MST and MAU teams – without your help this project would not have been possible. I would like to extend my gratitude to my supervisors, Dr Stephen Butler and Professor Steve Pilling, for their valuable advice and support, also Dr John King for his advice on statistics, Rachel Haley for all her assistance with START queries, as well as the team of START research assistants, the researchers who responded to my requests for copies of their work, and the staff in the Research Department of Clinical, Health and Educational Psychology at UCL. I am grateful to my clinical supervisors for their support and flexibility which has allowed me, among other things, to travel around the country to conduct the interviews!

Finally, I would like to extend huge thanks to my family, particularly my parents and my partner, and friends, including my fellow trainees, for all their support.

Part 1: Literature Review

What is the appropriate structure and content of a service delivery fidelity measure for complex interventions for children and young people with conduct problems?

Abstract

Aims The purpose of this review is to establish what the existing literature and the measures identified in the review can tell us about the appropriate structure and content of a fidelity measure to evaluate the effective delivery of complex interventions to children with conduct disorder and associated psychosocial problems. This review includes systematic searches identifying existing adherence and fidelity measures for complex interventions for conduct problems; it will also examine measures and other relevant papers existing in the wider literature.

Method This paper details a literature search aimed at identifying existing adherence and fidelity measures used in complex interventions for children and young people with conduct problems. Due to limited data on such measures for conduct problems and child problems more generally, search terms were widened to also include measures used in complex interventions in adult mental health. A narrative synthesis approach was drawn upon in order to generate a useful understanding in an area where there is a disparate body of literature, and where more conventional methods might not offer the same insights.

Results 35 papers were included in the review. There is a body of literature on complex interventions for children and adolescents, some of which addresses adherence or fidelity. However, this was found to be largely at the level of treatment adherence and individual sessions as opposed to aiming to encapsulate the broader organisational elements of an intervention and the infrastructure surrounding it. Some of the literature offers conceptual frameworks, while some authors identify key

elements of interventions, but do not propose these in the form of a replicable or standardised framework. Assertive community treatment for severe and enduring mental health problems in adults is one area which gives more substantial attention to broader conceptualisations of essential service delivery characteristics that support interventions.

Conclusions This review concludes that there is scope for further work in the area, in developing a measure which is broader in terms of the domains addressed, offering a means of examining service delivery aspects of complex interventions, and which can be used across therapies, rather than being specific to one particular intervention.

Introduction

Conduct disorder is defined by the World Health Organization (WHO)'s International Classification of Diseases (ICD-10; WHO, 2007) as 'a repetitive and persistent pattern of dissocial, aggressive, or defiant conduct'. ICD-10 also requires for a diagnosis that such behaviour be present for six months or longer, and makes a clear distinction between conduct disorder and 'ordinary childish mischief or adolescent rebelliousness'. A report by the Office for National Statistics (2000) found a prevalence rate of 5.3 per cent for conduct disorder among children (boys and girls) aged five to 15 years (Meltzer et al., 2000). Conduct problems and associated behavioural difficulties exhibited by children and adolescents are significant social problems, costly in both personal and economic terms (see e.g. Foster and Jones, 2005; National Institute for Health and Clinical Excellence, 2006; Scott, Knapp, Henderson and Maughan, 2001).

In an effort to provide effective services for children and families affected by conduct disorder, and its precursor, oppositional defiant disorder, a number of complex interventions have been developed over the last 25 years or so. Complex interventions are by definition not straightforward to describe; the Medical Research Council defines them as comprising several interacting components, and exhibiting complexity in terms of one or several dimensions (including number and difficulty of behaviours, range of groups targeted, number and variability of outcomes, and degree of flexibility required in delivery) (Medical Research Council, 2008). These complex interventions for children include multisystemic therapy (MST; Henggeler, 1999; Henggeler and Borduin, 1990), as well as functional family therapy (FFT; Alexander and Parsons, 1982), multi-dimensional family therapy (MDFT; Liddle, 2010), Parent Management Training (PMT; Feldman and Kazdin, 1995) and Problem

Solving Skills Training (PSST; Kazdin et al., 1987). As interventions have developed, so too has the field of implementation science, and an understanding that dissemination of an intervention to settings other than that in which it was developed, or 'transport', requires a systematic approach. Schoenwald (2008) describes the development of an effective evidence-based approach to transporting MST to settings outside the research environment in which it was developed and validated. She describes how this approach sought to consider four levels of influence on implementation– youth/family, clinician, organisation, and service system. The latter two are of particular interest in this review, and Schoenwald identifies effective collaboration with systems and provider organisations, which creates a good fit between financial and organisational policies and MST, as essential in ensuring fidelity of implementation. While the investigation of fidelity in general is not comprehensive, Schoenwald and colleagues have produced a robust body of literature on the fidelity and transport of MST (Schoenwald, Sheidow and Letourneau, 2004) and effective transport of interventions (Schoenwald and Hoagwood, 2001; Schoenwald, Chapman, Sheidow and Carter, 2009). Fidelity and transport, while not equivalent, are overlapping concepts and this should be borne in mind in reviewing the literature.

It might be hypothesised that those features of the organisation and service system Schoenwald discusses are more than simply helpful in ensuring that new interventions can be implemented and done so with fidelity, perhaps they are in fact integral to the success of the therapy, essential rather than preferable. MST, which is arguably one of the most well developed intervention programmes for children and young people in terms of dissemination, has been demonstrated to be highly efficacious in the United States (e.g. Borduin et al., 1995; Schaeffer and Borduin,

2005), and when transported to some other countries, including Norway (Ogden and Halliday-Boykins, 2004; Ogden and Amlund Hagen, 2006) and New Zealand (Curtis, Ronan, Heiblum and Crellin, 2009). However, studies in Canada and Sweden (Leschied and Cunningham, 2002; Sundell et al., 2008) have shown less promising results. Sundell et al. (2008) did not replicate the success of MST seen in American and Norwegian studies in their Swedish sample, finding no statistically significant differences between control and experimental groups in terms of reduction of problem behaviours, improvement in family relations or improvement of social skills. Consequently, it is worth considering what factors may be related to the varying levels of treatment success when MST is implemented in different regions. One argument is that usual services in those countries which have shown more ambiguous results are less iatrogenic than those in the US, or at least offer a more robust service (Littell, Campbell, Green and Toews, 2005). The presence of the structures around MST that were also present in many studies of effectiveness might suggest different factors at play in how successfully the intervention is transported and how positive the outcomes are. Sundell et al. (2008) comment on the child welfare approach to young offenders in both Sweden and Norway which means services which are largely home-based are more common, and would not therefore be the preserve of MST in these two countries. While this might explain the difference between findings in Sweden and the US, it does not account for the difference between Sweden and Norway. However, in explaining their findings Sundell et al. cite the presence of national support for the project in Norway, compared with only local frameworks in Sweden, and the higher rate of residential placements in Norway, disadvantaging the Norwegian MAU group, as residential placements are considered to come with higher risks of an iatrogenic effect (see e.g.

Dodge, Dishion and Lansford, 2006). This highlights the importance of having measures available which are able to capture not only treatment adherence but the wider components of the system, that is to say the fidelity with which other elements of an intervention are transported, and the service delivery frameworks which support interventions. This question also arises in the literature on home treatment for adult mental health problems, where differences between usual services (with North American services considered poorer than those in Europe) are argued by some to explain the advantage displayed by North American services over their European counterparts (see Burns et al., 2002). A related issue, raised by Curtis, Ronan and Borduin (2004), is that the MST efficacy studies – those where one of the original developers of the intervention is directly involved in a consultancy or supervisory capacity – consistently demonstrate far greater effect sizes than do effectiveness studies – studies in which MST is evaluated in naturalistic settings to which it has been transported and where MST developers are usually absent.

The presence of a developer was initially considered the explanation for the difference between efficacy and effectiveness studies, in that their greater involvement leads to better adherence to the MST model when treatment is delivered. However, Curtis et al. (2009) have since added that the different comparison conditions are also likely to have been a factor here. An alternative explanation, and one which warrants more extensive consideration, is that particular systemic and organisational features which are supported by the structure of MST, but which might not be explicitly identified as essential components of it, are more variable when the intervention is transported, including in effectiveness studies. A related issue is that some management as usual (MAU) services may well contain the essential components of infrastructure of MST but not the associated formal MST

processes. These essential components might include quality of supervision, team ethos regarding how to achieve treatment aims and outcomes, consistency of approach to formulation across therapists, and supervisory arrangements and style.

In order to set the context for this review, it is useful to make some distinctions between the various terms that are used in this literature to describe how, and how well, interventions are implemented and delivered: *competence*; *adherence*; *fidelity*; and *dissemination*. While the terms are used differently in different arenas, the following definitions, derived from consideration of the literature, detail how the terms will be used for the purpose of this review. Competence describes one's level of skill in delivering a given intervention. Adherence means the degree to which a therapist (or supervisor or other person involved in delivery of a therapy) adheres to the manualised treatment procedures required by a particular model.

Ogden, Amlund Hagen, Askeland and Christensen (2009) highlight the useful distinction between programme fidelity; ensuring that the necessary elements for an intervention to be delivered are in place, and treatment fidelity; the core content of an intervention, and its exposure and 'dosage'. Programme fidelity, an organisational concept, could be said to describe the degree to which an intervention is delivered as intended *at all levels*, examining domains such as supervision, team structure, team ethos, communication style and patterns, and the approach to formulation used in teams. It should be noted that the term is sometimes used in the literature to refer to what might in fact be considered treatment fidelity.

While it is useful to keep these distinctions in mind, of equal importance is the recognition that the terms are at times used interchangeably, and that the definitions continue to evolve as the conceptualisation of the underlying principles so does.

We might consider successful dissemination to be a consequence of successful fidelity to a given model of complex intervention. Dissemination and fidelity can be best understood as related but not equivalent concepts; successful dissemination is not possible without first having means of establishing the level of programme fidelity. A successful intervention, in terms of achieving good outcomes as measured by established measures of agreed targets, is not the same as successful implementation or transport, that is to say how well an intervention adheres to the treatment principles and manual once transported. However, successful intervention is likely to be highly correlated with fidelity of intervention implementation.

This review aims to capture existing understanding of the service delivery components which are integral to successful complex interventions, and consider how these are measured, or might be. Though not unimportant, to examine the quality of these measures, in terms of their psychometric properties and route to development, is beyond the scope of this review, but would be a helpful extension of this work. Investigation into the wider question of service delivery factors has to date been largely absent from research in the area; no measure to effectively delineate these components and their relationship with transport in complex interventions for children and young people is currently available. This is in contrast with, for example, measurement of treatment adherence such as that exemplified by the established measures used routinely in MST which examine therapist adherence (Therapist Adherence Measure; TAM; Henggeler and Borduin, 1992), supervisor adherence (Supervisor Adherence Measure; SAM; Schoenwald, Henggeler and Edwards, 1998) and consultant adherence (Consultant Adherence Measure; CAM; Schoenwald, 1998) (see Appendices A-C for all measures). While research has looked at the effectiveness of the range of interventions aimed at children with

behavioural problems and their families, and MST demonstrates very good adherence practices, the degree to which fidelity is more widely considered in the literature is variable. A little over twenty years ago Moncher and Prinz (1991) reported in their evaluation of treatment outcome studies that fifty-five per cent ignored fidelity. We might understand the term *fidelity* as it is used in this instance as referring broadly to any strategies for ensuring an intervention is delivered as intended, at any level. Borrelli and colleagues (Bellg et al., 2004; Borrelli et al., 2005) sought to update and expand the literature, developing a framework within which to enhance fidelity and later using this framework to develop a treatment fidelity measure with which they evaluated articles published over a ten-year period in five major journals. They found a mean proportion adherence to treatment fidelity strategies, relating to design, training, delivery, receipt and enactment, of .55; an improvement on Moncher and Prinz's finding of just 10 per cent. In order to answer our question, specific elements of practice, the content and structure of existing measures of fidelity, and knowledge of implementation science literature will be drawn upon.

In understanding the appropriate structure for a measure of complex interventions we must consider the following:

- Content and focus
- Structure
- Conceptual framework within which the measure positions itself

Our interest is in understanding whether we can measure services in a replicable way, characterising services in terms of the service delivery frameworks which support them. It might be that services which are successful have in common

particular service delivery frameworks which are more important, yet perhaps more opaque, than the specific details of the interventions offered by them. The aim here is to take programme fidelity a step further and discover how we might assess fidelity not to specific interventions or strategies, or even to principles which characterise one specific intervention, but rather to a set of overarching organisational and philosophical principles underpinning a range of interventions – which might look different in terms of what form 'therapy' takes – known to be successful for a specific difficulty, namely conduct problems.

Waltz, Addis, Koerner, and Jacobson (1993) discuss the methods used at the time they were writing, nearly twenty years ago, to collect fidelity data – manipulation checks, tests of treatment integrity, and assessment of therapist competence and adherence – and make recommendations for future developments, including all aspects of therapist competence being defined relative to the treatment manual being used, carefully fitting manipulation checks to the questions asked, and employing adherence measures which assess the degree to which therapy includes behaviours which are unique and essential, essential but not unique, acceptable but not necessary, and proscribed. These recommendations have come to fruition in many pieces of research but, as we will see, by no means all. Harachi, Abbott, Catalano, Haggerty, and Fleming (1999) highlighted the need to move away from what has become known as the ‘black box’ approach to implementation, that is to say, one which lacks transparency into what interventions are doing, and towards one which allows elaboration of the specific mechanisms through which change occurs. This review proceeds with this in mind.

Method

Literature search strategy

A systematic search of the literature was conducted between November 2010 and February 2011 using computerised databases (MEDLINE, PsycINFO, Cochrane Library). Searches were limited to articles in English or those for which an English translation was available. Combinations of the following keywords were used: *adherence, fidelity, competence, implementation, measure, scale, tool, framework, index, checklist, child, adolescent*. The term *youth* was also considered but not included as initial searches showed that it mapped to the subject heading *adolescent*. These terms were derived from a review of key articles and a scoping search of available online literature. The term *integrity* is sometimes used interchangeably with *fidelity*; however exploratory searches again indicated that excluding this term from the main search strategy did not result in the loss of relevant articles, and its inclusion did not increase the number of unique articles. Additional terms were subsequently added in to narrow the focus of the search in the first instance; these included *conduct disorder, oppositional defiant disorder,* and *anti-social behavio**. Further searches were run which included specific names of interventions known to be used with young people with conduct disorder or behavioural problems more generally; these were *brief strategic family therapy, multidimensional family therapy and multisystemic therapy*. Exploratory searches indicated that to narrow the searches by including specific reference to *service delivery* was not productive, and it was considered preferable to use more inclusive terms, despite this creating a higher volume of results to visually check. The search was then run using terms which aimed to broaden it, thus capturing literature pertaining to psychotherapy more generally; these included *mental health, psychotherap*, intervention and treatment*.

In addition, articles or measures were included which were known to the author and her supervisors but which might not have been identified through the literature search, or which were identified from inspection of the reference lists of studies identified in the search process.

The total number of articles identified in the database search as possible for review was 2114 from the child literature and a further 1495 from the adult literature. Restricting the child literature to papers from the year 2000 onwards reduced the number from this area to 1764; the specific area under examination might be said to be a relatively new area of research so it was considered legitimate to limit the searches in this manner. From this, 28 were identified as being relevant to the review (meeting the criteria described below); a further 7 were included which were identified elsewhere as described above.

Paper inclusion criteria

Papers were included on the basis that they met at least one of the following criteria: concerned with implementing interventions for children with behavioural difficulties; presented frameworks for understanding the necessary factors for effective interventions; described fidelity measures which were concerned with service delivery features of interventions; concerned with issues of fidelity measurement. A number of papers related to organisational climate were identified, however these were excluded as it was considered that they would not directly inform the research questions we have set out to answer. Articles which were concerned with treatment adherence only, in adult interventions, were excluded.

Obtaining measures

An issue which arose during the early stages of reviewing the literature was that a number of articles cited measures used to assess fidelity and related constructs,

but did not append the measure. It was therefore necessary to contact authors directly to request copies of measures which were considered relevant to the search. Contact was made by email and all authors responded to the request positively, sending the measures and manuals, and in some cases including related articles. The following measures were obtained using this method: Leader Observation Tool (LOT; Eames et al., 2009); Fidelity of Implementation Rating Scale (FIMP; Knutson, Forgatch, Rains and Sigmarsdóttir, 2009); Therapist Behavior Rating Scale-Competence (TBRS 2; Hogue, Liddle, Singer, and Leckrone, 2005); Chicago Parent Program Fidelity Checklist (Breitenstein et al., 2010); Children's Psychosocial Rehabilitation Treatment Adherence Measure (CTAM; Williams, Oberst, Campbell and Lancaster, 2011); Borrelli Fidelity Framework (Borrelli et al., 2005); Semi-Structured Interview Protocol for Clinical Staff and Semi-Structured Interview Protocol for Agency Lead from the Pilot Study of Barriers and Facilitators to Functional Family Therapy (FFT) Implementation in NY State (Zazzali et al., 2008).

The Substance Abuse and Mental Health Services Administration (SAMHSA) measures described below were obtained via an online search, having first been identified through preliminary reading in the area, and not through the formal searches.

Reporting findings

In answering in as rich a way as possible the questions posed, the concept of narrative synthesis (see Popay et al., 2006) was drawn upon in collating and presenting the findings, in which the guiding principle is to generate a story which represents the key themes and understandings arising from examination of the literature, and which can be further developed as new literature is examined. As the literature on measurable means of evaluating the service delivery frameworks which

are common to complex interventions for children and young people is fragmented, and the question under examination not a more standard review question, a flexible approach to the review, utilising Popay et al.'s (2006) concept of narrative synthesis, was helpful.

Results

The articles in the review are divided into the following four categories: interventions for children with behavioural difficulties; conceptual frameworks or necessary factors for effective interventions; measures which address service delivery issues; and issues of measurement. 35 papers were deemed to meet criteria for inclusion, in that they were informative in understanding the appropriate content and structure for a measure of service delivery components of complex interventions for children and young people. Table 1 summarises the papers; in addition to the four categories above, and in order to help make sense of the findings, each was identified as informing the research question in terms of one or more of the following: 'structure', 'content' and 'other' (the third label to include those papers which were informative but which could not be said to directly relate to either structure or content).

The synthesis was guided by asking whether a given paper was concerned with issues of fidelity, and, further, at what level fidelity was explored, in those instances where it was. Moreover, examination of the literature generated an understanding that those papers concerned with key features of effective interventions would also be useful, helping to clarify what factors would be relevant to any measure which aims to assess service delivery elements of an intervention. Starting from the literature on implementation of complex interventions in the child

field, and moving gradually towards means of conceptualising effective interventions more broadly, and existing measures, in part taken from the adult field, a picture of what a measure of the service delivery aspects of complex interventions for children and young people with conduct problems might look like, started to emerge.

Table 1 Existing studies which can inform our understanding of the appropriate structure and content for a fidelity measure for the delivery of complex interventions to children and young people

Study	Area	Focus of paper	How does the study/measure inform our question as regards structure (S)/content (C)/other (O)?
Austin, Macgowan & Wagner 2005	Evidence-based practice (child)	Systematic review of 5 treatments for adolescent substance use	Fidelity checks in 20 % of sample (O)
Baer et al. 2007	Fidelity (adult)	Review of multi-site drug treatments	Identifies systemic factors influencing treatment (C)
Bellg et al. 2004	Fidelity (general)	Framework for measuring fidelity in health behaviour change interventions	Design, training, delivery, receipt, enactment (C)
Berzin et al. 2007	Fidelity (child)	Model fidelity in family group decision-making (FGDM)	Draws on a range of stakeholders' views (S)
Bond et al. 2000	Fidelity (adult)	Measuring fidelity in psychiatric rehabilitation	Multi-modal approach to collection of information (S)
Borrelli et al. 2005	Fidelity (general/issues of measurement)	Evaluation of fidelity in existing papers	25 item checklist (S); Design, training, delivery, receipt, enactment (C)
Breitenstein et al. 2010	Fidelity (child)	Establishing feasibility and validity of Fidelity Checklist for Chicago Parenting Program	One checklist comprising 2 scales (S); measuring adherence and competence (C)

Bruns et al. 2004	Fidelity (child)	Fidelity measure development	Interviews with multiple stakeholders, team observation measure, document review form, instrument to assess level of system support (S); experience of relevant stakeholders at different levels (including community and system) (C)
Durlak & DuPre 2008	Implementation	Factors affecting implementation	Framework for implementation factors – community level factors, provider characteristics, innovation characteristics, prevention delivery/support systems (C)
Eames et al. 2009	Fidelity (child)	Measure description	Session level factors only (C)
Fergusson, Stanley & Horwood 2009	Child intervention	Implementation	Absence of fidelity measure (O)
Fixsen et al. 2005	Fidelity (general)	Organisational fidelity in child interventions	Complexity of defining 'organisational fidelity' (O)
Forgatch, Patterson & DeGarmo 2005	Fidelity (child)	Application of rating scale to parent management training – the Oregon model (PMTO)	Session level factors only (C)
Garland et al. 2008	Evidence-based interventions (child)	Identifying common elements of psychosocial interventions for children with behavioural problems	Common factors across therapeutic modalities (C)
Gottfredson et al. 2006	Fidelity and effectiveness (child)	Experimental examination of fidelity and effectiveness across treatment approaches	Intervention specific fidelity measures employed (O)

Hogue et al. 2005; Hogue, Henderson et al. 2008; Hogue, Dauber et al. 2008	Fidelity (child)	Comparison of fidelity across approaches	Multidimensionality of problems (O)
Kling et al. 2010	Fidelity (child)	Comparison of parenting interventions	Information from parents and group leaders (C); session specific checklists (S)
Knutson et al. 2009	Fidelity (child)	Measure description	Intervention specific ratings (S;C)
Kumpfer & Alvarado 2003	Family interventions	Review of studies of parenting intervention	Principles of effective family interventions (C)
McGrew et al. 1994	Fidelity (adult)	Development of a fidelity scale (IFACT) for assertive community treatment (ACT)	3 subscales (S); staffing , organisation and service domains (C)
McHugo et al. 2007	Fidelity (adult)	Comparison of fidelity across approaches	Intervention specific fidelity scales (O)
Mueser et al. 2003	Fidelity (adult/child)	Description of scales	20 (Integrated Dual Disorders Treatment)/17 (Youth Integrated Community Treatment) items scored 1-5 (not implemented to fully implemented) (S); scores compared against protocol (C)
Ogden et al. 2009	Implementation (child)	Conceptual model of implementation components	Distinction between programme fidelity and treatment fidelity (O)
Scott et al. 2010	Implementation (child)	Description of study and discussion of role of fidelity	Emphasis on fidelity in dissemination (O)

SAMHSA 2003; 2009 (nb. 4 scales)	Fidelity (adult)	Measurement of a set of general operating characteristics of an organisation	12 items with each scoring 1-5 (not implemented to fully implemented) (S); systemic factors including programme philosophy, assessment, supervision, process monitoring (C)
Teague et al. 1998	Fidelity (adult)	Protocol for ACT fidelity scale (Dartmouth Assertive Community Treatment Scale)	3 domains of between 7 and 11 items, with each item scoring 1-5 (from not implemented to fully implemented) (S); domains on human resources, organisational boundaries, nature of services (C)
Turner & Sanders 2006	Dissemination (child)	Evaluation of evidence for specified parenting practices; discussion of evidence-based practice dissemination	Programme /resource development; quality training; promotion of practitioner self-efficacy; workplace support and supervision (C)
Wandersman et al. 2008	Dissemination/ Implementation (general)	Conceptual framework for relationships between systems in implementation	Framework comprises 3 systems: Prevention Synthesis and Translation (offering summarised information about interventions), Prevention Support (which provides training and other support to users in the field); Prevention Delivery (which implements innovations in practice) (C)
Williams et al. 2011	Adherence (child)	Development of an adherence measure for child psychiatric rehabilitation (CPSR)	35 item, 6 point scale with some items reverse scored (S); session specific (C)
Zazzali et al. 2008	Implementation of evidence-based practices (child)	Description of pilot study of functional family therapy (FFT), conceptual framework and interview schedules	2 semi-structured scales (S); clinical staff – background, implementation, organisational context, mechanisms of diffusion, overall assessment, future prospects; agency lead – adoption of new programs, program change (C)

Review findings

Interventions for children with behavioural difficulties. The synthesis starts with consideration of what the literature around interventions for children with conduct problems tells us broadly about the role of adherence and fidelity in this area. In line with the findings of Borrelli and colleagues described earlier (Bellg et al., 2004; Borrelli et al., 2005), this review finds that while the number of studies across the health literature generally which fail to consider fidelity is seemingly considerably smaller than it was 20 years ago, not all effectiveness studies looking at interventions for children include a fidelity or adherence check. A recent example in the child field is the study by Fergusson, Stanley and Horwood (2009) which provided preliminary data on the efficacy and cultural acceptability of the Incredible Years Basic Parent Programme in New Zealand. Their study incorporated measures of child behaviour and parent satisfaction; however the authors comment that the study lacks a measure of fidelity. We might argue that without fidelity measures any conclusions about the effectiveness or impact of an intervention are to be considered with caution, as we have no indication of the degree to which the intervention was representative of the intervention as intended or as compared with other incarnations of the programme.

A review of treatments for adolescent substance use by Austin, Macgowan and Wagner (2005) found fidelity checks in only one of the five therapies investigated (incidentally the one in which such a measure was in place was MST, and in this study adherence was shown to be a problem).

Recently a number of interventions in the child field have started to use adherence measures more routinely. These include the Chicago Parent Program (Breitenstein et al., 2010), Incredible Years (e.g. Eames et al., 2009; Scott et al.,

2010), Child Psychiatric Rehabilitation (Williams et al., 2011) and the Oregon model of Parent Management Training (Forgatch, Patterson and DeGarmo, 2005). The LOT (used by Incredible Years researchers in the UK), the Chicago Parent Program Fidelity Checklist; the FIMP (Fidelity of Implementation Rating System; used in the Oregon Parent Management Training programme); the CTAM (Children's Psychosocial Rehabilitation Treatment Adherence Measure; used in child psychiatric rehabilitation); and Hogue et al.'s 2005 Therapist Behavior Rating Scale (used for both Multi Dimensional Family Therapy and individual CBT for adolescent drug abuse), offer a range of structures and styles, but have in common session level factors as their focus, considering for example how therapy sessions are conducted, what techniques are employed, and degree of engagement. Their concern is what happens in sessions and not by and large the organisational structure or service delivery frameworks around interventions.

Similarly, Berzin, Thomas and Cohen (2007) assessed model fidelity in Family Group Decision-Making (FGDM). A standardised measure of fidelity was not employed, rather a range of measures aimed at eliciting views of a range of stakeholders were used. The constructs examined relate to session by session interaction. Kling, Forster, Sundell, and Melin (2010) assessed treatment fidelity in a study which looked at Parent Management Training (PMT) with varying levels of therapist support. While this study did employ a number of checks for fidelity to the model and child outcomes, again there appears to be no measure of fidelity as regards the wider service delivery context. The Strengthening Washington DC Families Project (Gottfredson et al., 2006) looked at four family-based interventions for child antisocial behaviour and its precursors. This review employed an adherence measure specific to each intervention; it did not seek to characterise the measures

using a common framework, and the measures used were session specific and aimed at the level of intervention in individual cases.

This is a common finding; those pieces of research which do use fidelity checks do so at the session level and not at a level which illuminates the systemic or service delivery features of a given therapy or intervention; that is to say, adherence or treatment fidelity rather than programme fidelity are more commonly addressed.

Garland, Hawley, Brookman-Frazee and Hurlburt (2008) describe an example of efforts to characterise common elements in the context of an intervention for children with disruptive behaviour problems, within the wider context of understanding barriers to effective implementation of evidence-based practice. This was achieved through a review of eight interventions for this population and a modified Delphi technique to validate the commonality of the elements identified. Garland et al. comment that, as we have seen, much of the implementation and fidelity literature concentrates on individual treatment protocol implementation; here they present a 'complementary approach', which considers the value of taking an across-treatments approach to delivering and assessing interventions and their success. While this is pertinent in that it reviewed common elements across different interventions, it did not do this at a service delivery level; rather it identified session level elements such as affect education and modelling. However, while the content in this article is closely tied to the therapeutic features as they relate to the specific delivery of individual sessions, the principle of overlapping features which are common in different treatment protocols serving the same clinical population is interesting. The authors acknowledge that commonality across interventions does not in and of itself prove centrality in therapeutic success, but it is a helpful indicator. Further, Garland et al. make reference to the role of the 'meta-aspects' of

interventions which are not identified because they are not specific techniques, for example, having an overarching conceptual framework which ties elements of the intervention together. While the focus remains largely therapeutic rather than organisational, this fits with the premise that there are likely to be, in any given intervention, elements which are inherent but not explicitly prescribed. To build on Garland et al.'s meta-aspects, we might hypothesise that features could include, for example, the degree to which workers subscribe to a shared model of care, protocols around how a team interacts with other service providers, the organisational structure within which a team sits, the comprehensiveness of supervision, and the extent to which an assertive approach to engagement is adopted.

Kumpfer and Alvarado (2003) reviewed two national studies in the United States that explored the availability and features of effective preventative family interventions for children and adolescents exhibiting problem behaviours. The authors identified 13 principles of effective family-focussed interventions (see Appendix D); to summarise, this included offering comprehensive, multi-component interventions, with increased 'dosage' for families at highest risk, tailoring interventions to cultural traditions of families, and a collaborative process through which they are delivered. The features identified can be seen to represent elements of service delivery which are not specific to a particular model of therapy. What we see in this example is a description of key characteristics for effective interventions for young people with complex problems, but without a systematic framework for evaluating the degree to which different interventions might include them.

Hogue and colleagues (Hogue et al., 2005; Hogue, Henderson et al., 2008; Hogue, Dauber et al., 2008) are a group of researchers who have examined the relationship between adherence, competence, and differentiation, that is the degree to

which an intervention's main principles are theoretically distinctive, in the child field. They emphasise the importance of fidelity in this area, commenting that 'the complexity of delivering intensive, multi-component preventions to the highest-risk adolescents and families demands rigorous fidelity monitoring and evaluation to ensure successful model implementation and adaptation' (Hogue et al., 2005, p.207).

What we see from examining the above articles, both those investigating specific interventions and those considering the degree to which fidelity is considered in such trials, and considering the key themes emerging from them, is that issues of fidelity and its relationship with outcomes are not routinely considered in the child field, despite exceptions such as MST. Where it is considered, there are varying definitions of it. Fidelity might refer to adherence to a specific set of intervention strategies described in a manual, or to a commitment to implementing service development according to a particular set of principles. The measurement of it therefore might take various forms. We also see that the uniting feature in those examples that do measure fidelity is the focus on session level therapeutic features, and not service delivery elements.

Conceptual frameworks or necessary factors for effective interventions.

In terms of those articles identified by the search which were specifically concerned with characterising successful interventions, a number described not a measure or scale but a conceptual framework. These are taken from both the child and adult fields.

Some studies are concerned with narrowly measuring adherence to the model, while others look more broadly at factors in successful dissemination. The former do not tend to consider the broader organisational aspects which might vary across

services, while the latter generally offer concepts rather than working frameworks with which to measure the functioning of a team.

Zazzali et al. (2008) describe a conceptual framework, based on the diffusion of innovations and organisational behaviour literature, and on accounts of implementation of evidence-based practice (EBP) in mental health, which aims to address the multiple levels of analysis required to provide an integrated approach to understanding implementation. The framework posits that organisational facilitators (leadership, resources, culture and climate, and structure), drivers of adoption (rational/technical, resource dependencies and institutional effects) and characteristics of EBP (flexibility and feasibility) mutually influence the adoption of an EBP, and its subsequent implementation and continuance. From this framework, Zazzali et al. developed two semi-structured interview schedules (see Appendix E), the purpose of which was to understand factors related to the implementation of Functional Family Therapy in one area in the US, New York State. The first schedule, aimed at clinical staff, is made up of six sections (background information, process of implementation, organisational context, mechanisms of diffusion, overall assessment and future prospects of the programme, plus a one-question section for feedback on interview process and questions) comprising 30 questions, while the schedule for agency leads is divided into two sections (adoption of new programs and program change experiences, and adoption of FFT and program change experiences) and comprises 25 questions. Responses to the semi-structured interview are considered in the context of the conceptual framework in order to understand the particular service under scrutiny. While different questions are asked of agency leads and clinical staff, the process relies on information from both supervisory or managerial and clinical staff – recognising the different experiences of the two

groups yet still considering them in combination – in order to effectively characterise a service or implementation process using multiple informants.

Zazzali et al. (2008) offer an integrated approach in which the infrastructural features of successful interventions are considered, a conceptual framework is offered and from this, questions derived which aim to characterise the system in focus. What Zazzali et al.'s framework and semi-structured interviews do not seem to offer is a manualised system with which to assess the features in question. Further, it focusses on attitudes to, and experiences of, implementation, but not on the sustainment of specific organisational features once established.

Ogden et al. (2009) present a conceptual model of the components influencing the outcomes of evidence-based programmes for children and young people, considering the relationships between the following factors and success of implementation: programme development, dissemination, adoption, readiness/awareness, fidelity/adherence, implementation, adaptation, context, and outcomes. These might be considered to describe overarching principles which would inform the way in which a service delivery framework was established for any intervention. As with other articles which present conceptual models of the central factors in implementation, or indeed transport, the authors do not extend the framework to one in which fidelity to the essential factors might be measured, nor to the details of how any one given factor might be quantified. Ogden et al. usefully point out that fidelity is measured in vastly different ways in different contexts, for example, through site assessments before implementation, through adherence measures, or through observation of sessions, as well as commenting that factors which are extra to the core components of an intervention, for example therapist enthusiasm, might be at play and thus impact on outcomes. These issues are

important in considering the appropriate means of characterising and measuring services.

Turner and Sanders' (2006) conceptual model of dissemination of the Triple P parenting programme describes factors which influence successful dissemination, including organisational factors, for example, the availability of adequate supervision, line management support, and adequate funding for programme implementation. Turner and Sanders also speak in more general terms about their recommendations for the implementation of evidence-based interventions, identifying the following non-programme specific, organisational aspects as key: programme and resource development; quality training; promotion of practitioner self-efficacy; workplace support; and supervision. These are elements which are likely to be elaborated to differing extents across different intervention guidelines or manuals, but which will invariably impact upon success of an intervention in terms of how it is implemented and embedded in an organisational structure. These are central components which could be considered valuable elements of any measure which was aimed at characterising the service delivery frameworks supporting interventions. While Turner and Sanders highlight these key elements, they do not propose a standalone measure of fidelity or a framework which aims to characterise the nature of interventions which are successfully disseminated.

Similarly to Turner and Sanders, but this time in the adult substance misuse field, Baer et al. (2007) offer a review of multi-site drug treatments and identify key elements in their success: employment of treatment manuals; provision of standardised competency-based training; use of rating scales for adherence measurement and quality improvement; employment of specific performance and certification procedures; monitoring of intervention delivery via review of sessions

(in vivo or taped); supervision and support processes for those delivering treatments; and regular oversight of these supervision and support procedures. There is a bringing together of a number of systemic factors which are likely to impact upon the nature of a treatment model, the way in which it is delivered and the degree to which implementation is successful. Here again we see a helpful framework for understanding implementation success, but without a structure within which to assess the elements systematically.

Borrelli and colleagues (see e.g. Borrelli et al., 2005) describe a fidelity framework which has a particular focus on the implementation of interventions in the context of research. The measure itself (personal communication, 6 January 2011) consists of what are described as five categories of treatment fidelity; design, training, delivery, receipt and enactment. While the majority of the items constituting the measure are concerned with the specific details of delivery of interventions, the training section does consider the approach to the hiring of staff delivering interventions and standardisation of training. These ideas start to move closer to the question of how to measure implementation at a higher, more systems-oriented level.

Durlak and DuPre (2008) specifically address implementation; they conducted a meta-analysis from which they identified factors impacting on successful implementation of promotion and prevention programmes in the community, as part of their analysis of the implementation element of the diffusion process. The authors use an ecological framework to conceptualise the relationship between contextual factors and outcomes. Durlak and DuPre identify five areas related to outcomes: community level factors (e.g. funding and policy); provider characteristics (e.g. perceived need for innovation, skill proficiency); characteristics of the innovation (e.g. compatibility and adaptability); factors relevant to the

prevention delivery system (organisational capacity) (e.g. integration of new programming, shared vision, coordination with other agencies, formulation of tasks, leadership and managerial/supervisory/administrative support); and factors related to the prevention support system (e.g. training and technical assistance).

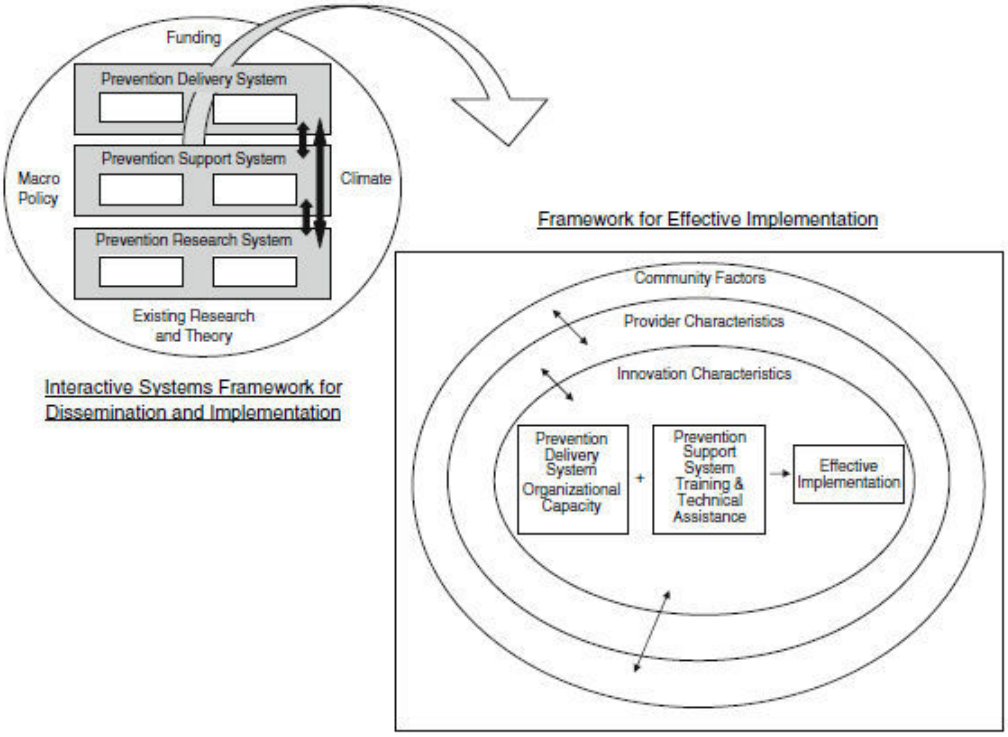

Figure 1 Durlak and DuPre's ecological framework for understanding effective implementation (Source: Durlak and DuPre, 2008).

Here the factors thought to relate to successful implementation and the manner in which they interact are delineated, but a framework by which these can be systematically evaluated across settings or therapies is not proposed. Durlak and DuPre also raise the interesting question of the degree to which fidelity is always paramount, or whether in fact adaptation or reinvention to suit local needs is also important. They assert that adaptation is a helpful and appropriate part of

implementation; it might be that adaptation, or the capacity for it, is in fact a feature which should be measured in characterising interventions.

Wandersman et al. (2008) propose the Interactive Systems Framework for Dissemination and Implementation (ISF). This is a framework for conceptualising relationships between systems in implementation, with particular reference to prevention interventions. The framework comprises three systems: the Prevention Synthesis and Translation System (offering summarised information about interventions); the Prevention Support System (which provides training and other support to users in the field); and the Prevention Delivery System (which implements innovations in practice). The framework is intended to be used by different types of stakeholders, for example, funding bodies, practitioners and researchers, as a means of understanding the requisite action from particular parts of a system to ensure effective implementation of an intervention. The authors comment that the ISF also highlights the need for communication between different stakeholders. While the ISF describes the relationships and roles of different system parts considered central to the implementation of a given intervention, it does not operationalise the means by which an organisation might measure its capacity to fulfil the needs of the population served by an intervention. Nevertheless, the principles it outlines might be valuable in understanding what a fidelity measure of service delivery factors might look like, at least in part.

We see that a number of authors highlight the value of conceptual frameworks through which to understand how interventions are best implemented, citing service delivery level factors which are likely to be relevant across different interventions. A logical progression of this thinking might be a measure which assesses in a quantitative, replicable way the degree to which any given service is

delivering interventions in a manner which supports successful implementation and fidelity to underlying principles and specific techniques, based on these conceptual frameworks. The work of Zazzali and colleagues (2008) and Borrelli and colleagues (2005) are examples of steps in this direction.

Measures which address service delivery. In order to make effective use of the review process in an area without an established evidence base, it was necessary to widen the search to include information taken from the adult mental health field. As searches of the adult field yield more items which are closer to the measure we are considering in our research question, we will begin here with the adult field, moving on to consider what exists in the child field.

One area in which there is fidelity measurement which takes a different approach, looking at service delivery aspects rather than either the narrower session by session elements or the more general conceptual frameworks, is in the work of researchers connected with Dartmouth College in the United States. The key measure emerging from this group is the Dartmouth Assertive Community Treatment Scale (DACTS; Teague, Bond and Drake, 1998). This measure, which examines assertive community treatment (ACT) for adults with severe and enduring mental health problems, could be said to bridge the gap between the narrowest and broadest attempts to characterise complex interventions in an effort to establish fidelity to the model and understand what key elements are necessary for effective transport, systematically evaluating three key domains: human resources; organisational boundaries; and nature of services. Additional scales which derive from or are similar to the DACTS are the Integrated Dual Disorders Treatment Fidelity Scale (Mueser, Noordsy, Drake and Fox, 2003; see also Wilson and Crisanti, 2009), the Illness Management and Recovery Fidelity Scale (Substance Abuse and Mental

Health Services Administration; SAMHSA, 2009), the Family Psychoeducation Fidelity Scale (SAMHSA, 2009), the Index of Fidelity of Assertive Community Treatment (IFACT; McGrew, Bond, Dietzen and Salyers, 1994) and the Supported Employment Fidelity Scale (SAMHSA, 2009).These scales are designed for use in conjunction with the General Organizational Index (SAMHSA, 2003). A further scale related to the DACTS and closest to a framework for assessing complex interventions for children is the Youth Integrated Community Treatment Fidelity Scale (adapted from the Integrated Dual Disorders Treatment Fidelity Scale; Mueser et al., 2003), which examines services for young people with co-occurring mental health and substance misuse disorders. These measures all address organisational level issues of fidelity in relation to one specific therapeutic intervention. It should be noted that the DACTS authors comment that where no ACT team exists within an agency, the measure may be used to evaluate non-ACT teams; however the focus of DACTS is clearly the ACT model, so it would presumably need a degree of modification if it were to be used in this way.

What is evident in all these measures is a concern with the manner in which teams operate, in terms of, for example, caseloads, leadership or supervision, where services are provided, and staff training. They concern themselves to varying degrees with more specific treatment fidelity issues too, but they all contain elements of measurement of programme fidelity, demonstrating an awareness in this area of health service research of the importance of this aspect of service provision.

Furthermore, the DACTS and the related measures demonstrate the value of multiple sources of evidence, including both verbal accounts and written sources, and information gathered from individuals working at different levels of an organisation.

This multi-informant approach increases validity and therefore increases the usefulness and acceptability of a measure.

McHugo et al.'s 2007 study is an example of a piece of research concerned with issues of fidelity in implementing evidence-based practice, again in the adult mental health field. It looked at five interventions – supported employment; family psychoeducation; illness management and recovery; integrated dual disorders treatment; and assertive community treatment – implemented in 53 sites across eight states in the US, and found varying levels of fidelity between interventions, with fidelity increasingly across all interventions over time. While they were looking at a range of similar interventions for people with severe mental health problems, and comparing fidelity across all sites, fidelity was nevertheless assessed with measures specific to each model, indicating that the focus, while more systemic, remained on specific details of a prescribed intervention, rather than with structural features which might be applicable to a range of different interventions aimed at the same population. The measures varied in terms of whether they were concerned with structure of practice or the clinical expertise required to deliver an intervention.

Returning to the child literature, Bruns, Burchard, Suter, Leverentz-Brady and Force (2004) report on the development, psychometric characteristics and utility of their Wraparound Fidelity Index (WFI; Suter, Burchard, Force, Bruns and Mehrtens, 2002) which assesses adherence to the wraparound model of intervention for children with emotional and behavioural difficulties and relies on multiple informant report. The index looks at the experience of relevant stakeholders throughout a treatment process, examining, for example, whether families feel they are active partners in the process, whether professionals demonstrate cultural competence, and the degree to which teams encourage involvement with activities in

the community. However, it does not address service delivery issues in terms of how the team is organised, rather its reference point is the treatment experience, as reported by clients. It does however tap into some themes which might be expected in a measure concerned with broader systemic issues, such as whether a family has been asked about their satisfaction with the service within the last three months and whether the team relies mostly on professional services, as compared with informal family or community support.

Fixsen, Naoom, Blase, Friedman and Wallace (2005) address in their synthesis of the literature on implementation research the issue of organisational level fidelity, making reference as we might expect to the DACTS. They also refer to the Child and Adolescent Functional Assessment Scale (CAFAS; see Hodges and Kim, 2000) used in Michigan in the United States, grouping this with other measures which assess organisational level issues. However, this scale is a measure that assesses the level of children's functioning from a clinical perspective, analysing the characteristics of children for whom treatments were most and least successful, identifying those problems which were most intractable in usual services, and from this identifying services considered good candidates for implementing evidence-based programmes. It seems that CAFAS is a means of identifying need based on clinical data used at a population level. While it is universal in the sense that it can be used across services, and as a tool for understanding where new evidence-based practices would be best placed, it is not in fact a measure which assesses programme fidelity as its inclusion in this section of their review might suggest.

Issues of measurement. Bond et al. (2000) discuss issues around measurement of fidelity in their article which looks specifically at measuring fidelity in psychiatric rehabilitation for adults with severe and enduring mental health

problems. A key point in relation to our question is the value of a multi-modal approach to collecting information which informs the outcome of the measure, that is to say, an approach to measurement which includes, for example, surveys of staff, chart reviews and observations of team meetings. It seems likely that the specific means of gathering information will vary depending on the focus of the measure and the nature of the team under examination, but a key idea is that of using multiple sources of information to generate as comprehensive and reliable a picture as possible of the constructs in question. A measure of this nature must be conceived with not only a consideration of what information it relies upon but how this information will be gathered, taking into account the fact that the source of information, and the means by which it was obtained, partially inform how we understand it.

Discussion

This review captures some key elements of the developing area of fidelity at a service delivery level, and puts this in the context of the wider fidelity field. A diverse selection of literature has informed the understanding developed from this review. The review suggests that there is scope for further exploration of the possibilities for employing fidelity scales which capture the service delivery elements of complex interventions and which might be used across services.

The role of fidelity as it relates to organisational, systemic or service delivery factors in efficacy and effectiveness research continues to evolve. The importance of adherence scales which identify the degree to which a therapist is delivering an intervention as prescribed by the relevant manual, at the level of individual sessions, though not to be taken for granted, has been increasingly recognised. The

development of the competence frameworks for systemic therapies (Pilling, Roth and Stratton, 2010) and child and adolescent mental health services (Roth, Calder and Pilling, 2011) are examples of the increasing importance placed upon standardised, measurable and demonstrably competent professional practice throughout services. What has hitherto been relatively unexplored is the way in which the service delivery features of a given team delivering an intervention might impact on the transport of the intervention and the efficacy of the work delivered. Relatively little is known about how to uniformly measure and characterise services in terms of their organisation. Whether services look similar at the organisational level, irrespective of the specific interventions offered, might be indicative of their relative success. That is to say, it might be a number of organisational factors which surround a given intervention which are intrinsic to its success, rather than, or as well as, the session level features of the intervention.

To address the review question more directly, examining the literature suggests that the appropriate structure and content for a fidelity measure of service delivery frameworks supporting complex interventions for children and young people with conduct disorder might be best informed by a combination of existing measures of fidelity in the adult field and understanding drawn from conceptual frameworks which look at necessary factors in effective interventions. Measures of this nature in the child literature tend to concentrate on the use of specific features of an intervention and the degree to which particular activities are present and indicative of strict adherence to the therapeutic model. Adult measures exist which take a more systemic or organisational approach, aiming to uncover the nature of a team in terms of its structure. These measures use multiple informants to improve the reliability of their outcomes.

What we have not seen is evidence of measures which offer a characterisation of teams which might offer different interventions from one another but to the same populations and often with the same underlying principles. Fidelity and issues of what supports successful implementation are largely addressed at the level of specific interventions (whether or not a measure is concerned with session level or service delivery features), and not at the level of a particular clinical population, nor the service context that supports an intervention, making comparison between interventions or teams in terms of their infrastructure, and the relationship between infrastructure and outcomes, difficult. As table 1 showed, there is important information that can be gleaned from existing measures and writing in the area. In terms of structure, existing measures tend to employ five point rating scales, use subsections to divide topics, and often rely on multiple informants, including both reported and observed data, which might be from staff or clients, as well as other stakeholders, and might be verbal or written. Conceptually, the measures and articles reviewed are diverse, however several areas emerge which are common across the papers: training, organisational support, monitoring systems, clarity of programme philosophy or ethos, assessment procedures, supervision arrangements, organisational context and, indeed, the use of treatment adherence measures. Though by no means exhaustive, this list describes items a service delivery measure might consider.

This review of the literature suggests that there is an opportunity to build on what exists in terms of measuring and understanding those factors which influence the degree to which interventions are successful. In order to understand the relationship between the various service delivery factors, scales which fill the gap between the conceptual frameworks we have seen described and session by session

adherence, might be developed. These might look specifically at the contextual factors in a systematic manner, drawing on the literature aimed at identifying those service delivery features of an intervention necessary for it to be successful. The ultimate goal might be to develop a pan-therapeutic measure which could discriminate between different types of services based on their infrastructure and elucidate the relationship between infrastructure and outcomes.

It should be noted that issues related to organisational climate overlap to some extent with issues around the ways in which services are organised. (For a comprehensive review of the relationship between organisational climate and outcomes in children's services see Glisson and Hemmelgarn, 1998). However, to widen our focus here would be to move away from the important focus on the relationship between service delivery and fidelity to guiding principles, and outcomes. In so doing we might run the risk of shifting our focus to the impact that organisational climate has on staff and the performance of a service, which, while important, is not in and of itself a central concern in this context.

The nebulous and shifting nature of the understanding of fidelity and implementation, and the multiple definitions surrounding them, make a comprehensive review of the question challenging. What this review has aimed to do is to bring together understanding of how fidelity measures look structurally and what their focus tends to be, as well as what literature on effectively implementing complex interventions for young people tells us, in order to better understand how a measure which takes a broader service delivery perspective might look.

Such a measure should draw on the knowledge existing in the field as regards those factors which are important in effective dissemination and implementation of interventions, and which we have seen in the articles reviewed here. Further, it

should utilise multiple informants and sources of information to ensure a comprehensive and representative overview of a given service. Effective, pan-model measurement of these service delivery level domains would be an interesting development in the pursuit of understanding the nature of services in terms not only of the models from which clinicians work, but also of the infrastructure which supports complex interventions and evidence-based practice.

It is important to comment on the possible limitations of this review, noting that the ideas it sets out to examine are not clearly established in the literature, meaning that a flexible approach has been necessary in bringing together the information gathered. There are likely to be alternative approaches to establishing the appropriate structure and context for a measure of service delivery, which might have been used. This approach has attempted to establish what already exists and might have been unhelpfully narrow as a result. Where the focus has been on existing measures, perhaps it might have been more productive to not be overly concerned with measures per se, but rather to consider in greater detail the fundamental elements of service delivery which are specified in the range of complex interventions currently employed most widely in offering interventions to this population, and use this as a basis from which to explore the area.

Further, lack of clarity around the terms *adherence* and *fidelity* abounds. While the decision was made to be relatively flexible in our understanding of them, perhaps in so doing we are perpetuating the difficulties around their use, where in fact to deconstruct the terms themselves might be to illuminate our understanding of this growing and yet at times opaque field.

Moreover, literature from the adult field has been drawn on to inform our understanding, and with little reference to the differences between child and adult

services. While it might be arguable that the distinction is less important at the service delivery level, the difference should nonetheless be borne in mind. Evidence-based practice as applied to children is necessarily different in terms of the centrality of the developmental focus and importance of the family context, as well as in terms of practical issues around where services are provided and by whom (Hoagwood, Burns, Kiser, Ringeisen and Schoenwald, 2001). These differences mean that any attempt to transpose adult tools onto child interventions and services is likely to be a less than perfect fit, and requires reflection on those elements which make them different.

The above shortcomings notwithstanding, the potential for exploration of this area is apparent, and offers opportunities to think about the way in which we characterise services afresh, in a manner which has meaning and application in modern health and social care services.

References

Alexander, J.F. and Parsons, B.V. (1982) *Functional Family Therapy: Principles and procedures*. Carmel, California: Brooks and Cole.

Austin, A.M., Macgowan, M.J. and Wagner, E.F. (2005) Effective Family-Based Interventions for Adolescents with Substance Use Problems: A Systematic Review. *Research on Social Work Practice, 15*, 2, pp.67-83.

Baer, J.S., Ball, S.A., Campbell, B.K., Miele, G.M., Schoener, E.P. and Tracy, K. (2007) Training and fidelity monitoring of behavioral interventions in multi-site addictions research. *Drug and Alcohol Dependence, 87*, pp.107-118.

Bellg, A.J, Borrelli, B., Resnick, B., Hecht, J., Sharp Minicucci, D., Ory, M., Ogedegbe, G., Orwig, D., Ernst, D. and Czajkowski, S. (2004) Enhancing Treatment Fidelity in Health Behavior Change Studies: Best Practices and Recommendations From the NIH Behavior Change Consortium. *Health Psychology, 23*, 5, pp.443-451.

Berzin, S.C., Thomas, K.L. and Cohen, E. (2007) Assessing Model Fidelity in Two Family Group Decision-Making Programs. *Journal of Social Service Research, 34*, 2, pp.55-71.

Bond, G.R., Evans, L., Salyers, M.P., Williams, J. and Kim, H.-W. (2000) Measurement of Fidelity in Psychiatric Rehabilitation. *Mental Health Services Research, 2*, 2, pp.75-87.

Borduin, C.M., Mann, B.J., Cone, L.T., Henggeler, S.W., Fucci, B.R., Blaske, D.M. and Williams, R.A. (1995) Multisystemic Treatment of Serious Juvenile Offenders: Long-Term Prevention of Criminality and Violence. *Journal of Consulting and Clinical Psychology, 63*, 4, pp.569-578.

Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A.J., Czajkowski, S., Breger, R., DeFrancesco, C., Levesque, C., Sharp, D.L., Ogedegbe, G., Resnick, B. and Orwig, D. (2005) A New Tool to Assess Treatment Fidelity and Evaluation of Treatment Fidelity Across 10 Years of Health Behavior Research. *Journal of Consulting and Clinical Psychology, 73*, 5, pp.852-860.

Breitenstein, S.M., Fogg, L., Garvey, C., Hill, C., Resnick, B. and Gross, D. (2010) Measuring Implementation Fidelity in a Community-Based Parenting Intervention. *Nursing Research, 59*, 3, pp.158-165.

Bruns, E.J., Burchard, J.D., Suter, J.C., Leverentz-Brady, K. and Force, M.M. (2004) Assessing Fidelity to a Community-Based Treatment for Youth: The Wraparound Fidelity Index. *Journal of Emotional and Behavioral Disorder, 12*, 2, pp.79-89.

Burns, T., Catty, J., Watt, H., Wright, C., Knapp, M. and Henderson, J. (2002) International differences for home treatment for mental health problems. *British Journal of Psychiatry, 181*, pp.375-382.

Curtis, N.M., Ronan, K.R., and Borduin, C.M. (2004) Multisystemic treatment: A meta-analysis of outcome studies. *Journal of Family Psychology, 18,* pp.411–419.

Curtis, N.M., Ronan, K.R., Heiblum, N. and Crellin, K. (2009) Dissemination and Effectiveness of Multisystemic Treatment in New Zealand: A Benchmarking Study. *Journal of Family Psychology 23*, 2, pp.119-129.

Dodge, K.A., Dishion, T.J., and Lansford, J.E. (2006) *Deviant peer influences in programs for youth: Problems and solutions.* New York: Guilford Press.

Durlak, J.A. and DuPre, E.P. (2008) Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology, 41*, pp.327-350.

Eames, C., Daley, D., Hutchings, J., Whitaker, C.J., Jones, K., Hughes, J.C. and Bywater, T. (2009) Treatment fidelity as a predictor of behaviour change in parents attending group-based parent training. *Child: care, health and development, 35*, 5, pp.603-612.

Feldman, J. and Kazdin, A.E. (1995) Parent management training for oppositional and conduct problem children. *The Clinical Psychologist, 48*, 4, pp.3-5.

Fergusson, D., Stanley, L. and 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, 43*, pp.76-79.

Fixsen, D.L., Naoom, S.F., Blase, K.A., Friedman, R.M. and Wallace, F. (2005) *Implementation Research: A Synthesis of the Literature* (FMHI Publication number

231). Tampa, Florida: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Forgatch, M.S., Patterson, G.R. and DeGarmo, D.S. (2005) Evaluating Fidelity: Predictive Validity for a Measure of Competent Adherence to the Oregon Model of Parent Management Training. *Behaviour Therapy, 36*, pp.3-13.

Foster, E.M. and Jones, D.E. (2005) The High Costs of Aggression: Public Expenditures Resulting From Conduct Disorder. *American Journal of Public Health, 95*, 10, pp. 1767-1772.

Garland, A.F., Hawley, K.M., Brookman-Frazee, L. and Hurlburt, M.S. (2008) Identifying Common Elements of Evidence-Based Psychosocial Treatments for Children's Disruptive Behavior Problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 5, pp.505-514.

Glisson, C. and Hemmelgarn, A. (1998) The Effects of Organizational Climate and Interorganizational Coordination on the Quality and Outcomes of Children's Service Systems. *Child Abuse and Neglect, 22*, 5, pp.401-421.

Gottfredson, D., Kumpfer, K., Polizzi-Fox, D., Wilson, D., Puryear, V., Beatty, P. and Vilmenay, M. (2006) The Strengthening Washington D.C. Families Project: A Randomized Effectiveness Trial of Family-Based Prevention. *Prevention Science, 7*, 1, pp.57-74.

Harachi, T.W., Abbott, R.D., Catalano, R.F., Haggerty, K.P. and Fleming, C.B. (1999) Opening the Black Box: Using Process Evaluation Measures to Assess Implementation and Theory Building. *American Journal of Community Psychology, 27*, 5, pp.711-731.

Henggeler, S.W. (1999) Multisystemic Therapy: An Overview of Clinical Procedures, Outcomes and Policy Implications. *Child Psychology and Psychiatry Review, 4*, 1, pp.2-10.

Henggeler, S.W. and Borduin, C.M. (1990) *Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents*. Pacific Grove, California: Brooks/Cole Publishing Company.

Henggeler, S.W. and Borduin, C.M. (1992) *Multisystemic Therapy Adherence Scales* [Unpublished instrument]. Charleston, South Carolina: Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina.

Hoagwood, K., Burns, B.J., Kiser, L., Ringeisen, H. and Schoenwald, S.K. (2001) Evidence-Based Practice in Child and Adolescent Mental Health Services. *Psychiatric Services, 52*, 9, pp.1179-1189.

Hodges, K. and Kim, C.-S. (2000) Psychometric Study of the Child and Adolescent Functional Assessment Scale: Prediction of Contact with the Law and Poor School Attendance. *Journal of Abnormal Child Psychology, 28*, 3, pp.287-297.

Hogue, A., Dauber, S., Chinchilla, P., Fried, A., Henderson, C., Inclan, J., Reiner, R.H. and Liddle, H.A. (2008) Assessing fidelity in individual and family therapy for adolescent substance abuse. *Journal of Substance Abuse Treatment, 35*, pp.137-147.

Hogue, A., Henderson, C.E., Dauber, S., Barajas, P.C., Fried, A. and Liddle, H.A. (2008) Treatment Adherence, Competence, and Outcome in Individual and Family Therapy for Adolescent Behavior Problems. *Journal of Consulting and Clinical Psychology, 76*, 4, pp.544-555.

Hogue, A., Liddle, H.A., Singer, A. and Leckrone, J. (2005) Intervention Fidelity in Family-Based Prevention Counseling for Adolescent Problem Behaviors. *Journal of Community Psychology, 33*, 2, pp.191-211.

Kazdin, A.E., Esveldt-Dawson, K., French, N.H., Unis, A.S. (1987) Problem-solving skills training and relationship therapy in the treatment of antisocial child behaviour. *Journal of Consulting and Clinical Psychology, 55*, 1, pp.76-85.

Kling, A., Forster, M., Sundell, K. and Melin, L. (2010) A Randomized Controlled Effectiveness Trial of Parent Management Training with Varying Degrees of Therapist Support. *Behavior Therapy, 41*, pp.530-542.

Knutson, N.M, Forgatch, M.S., Rains, L.A. and Sigmarsdóttir, M. (2009) *Fidelity of Implementation Rating System (FIMP): The Manual for PMTO™*. Eugene, Oregon: Implementation Sciences International, Inc.

Kumpfer, K.L. and Alvarado, R. (2003) Family-Strengthening Approaches for the prevention of Youth Problem Behaviors. *American Psychologist, 58*, 6-7, pp.457-465.

Leschied, A.W. and Cunningham, A. (2002) *Seeking effective interventions for young offenders: Interim results of a four-year randomized study of Multisystemic Therapy in Ontario, Canada* [online]. Available at: http://www.lfcc.on.ca/seeking.html . Accessed 20 September 2010.

Liddle, H.A. (2010) Multidimensional Family Therapy: A Science-Based Treatment System . *The Australian and New Zealand Journal of Family Therapy, 133*, 2, pp.133-148.

Littell, J.H., Campbell, M., Green, S. and Toews B. (2005) Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17. *Cochrane Database of Systematic Reviews 2005*, Issue 4. Art. No.: CD004797. DOI: 10.1002/14651858.CD004797.pub4.

McGrew, J.H., Bond, G.R., Dietzen, L. and Salyers, M. (1994) Measuring the Fidelity of Implementation of a Mental Health Program Model. *Journal of Consulting and Clinical Psychology, 62*, 4, pp.670-678.

McHugo, G.J., Drake, R.E., Whitley, R., Bond, G.R, Campbell, K., Rapp, C.A., Goldman, H.H., Lutz, W.J. and Finnerty, M.T. (2007) Fidelity Outcomes in the

National Implementing Evidence-Based Practices Project. *Psychiatric Services, 58*, 10, pp.1279-1284.

Medical Research Council (2008) *Developing and evaluating complex interventions: new guidance* [online]. Available at: www.mrc.ac.uk/complexinterventionsguidance . Accessed 11 June 2012.

Meltzer, H, Gatward, R., Goodman, R. and Ford, T. (2000) *The mental health of children and adolescents in Great Britain. The report of a survey carried out in 1999 by Social Survey Division of the Office for National Statistics on behalf of the Department of Health, the Scottish Health Executive and the National Assembly for Wales.* London: The Stationery Office.

Moncher, F.J. and Prinz, R.J. (1991) Treatment Fidelity in Outcome Studies. *Clinical Psychology Review, 11*, pp.247-266.

Mueser, K. T., Noordsy, D. L., Drake, R. E., and Fox, L. (2003) *Integrated treatment for dual disorders: A guide to effective practice*. New York: Guilford Press.

National Institute for Health and Clinical Excellence (2006) *Parent-training/education programmes in the management of children with conduct disorders* [TA102]. London: National Institute for Health and Clinical Excellence.

Ogden, T. and Amlund Hagen, K. (2006) Multisystemic Treatment of Serious

Behaviour Problems in Youth: Sustainability of Effectiveness Two Years after Intake. *Child and Adolescent Mental Health, 11*, 3, pp.142-149.

Ogden, T., Amlund Hagen, K., Askeland, E. and Christensen, B. (2009) Implementing and Evaluating Evidence-Based Treatments of Conduct Problems in Children and Youth in Norway. *Research on Social Work Practice, 19*, 5, pp.582-591.

Ogden, T. and Halliday-Boykins, C.A. (2004) Multisystemic Treatment of Antisocial Adolescents in Norway: Replication of Clinical Outcomes Outside of the US. *Child and Adolescent Mental Health, 9*, 2, pp.77-83.

Pilling, S., Roth, A.D. and Stratton, P. (2010) *The competences required to deliver effective Systemic Therapies* [online]. Available at: http://www.ucl.ac.uk/clinical-psychology/CORE/systemic_framework.htm. Accessed 22 May 2012.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K. and Duffy, S. (2006) *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A product from the ESRC Methods Programme.* Lancaster: Institute of Health Research.

Roth, A.D., Calder, F. and Pilling, S. (2011) *A competence framework for Child and Adolescent Mental Health* Services [online]. Available at: http://www.ucl.ac.uk/clinical-psychology/CORE/child-adolescent-

competences/CAMHS%20Competences%20Framework_V1%20(2).pdf . Accessed 18 June 2012.

Schaeffer, C.M. and Borduin, C.M. (2005) Long-Term Follow-Up to a Randomized Clinical Trial of Multisystemic Therapy with Serious and Violent Juvenile Offenders. *Journal of Consulting and Clinical Psychology, 73*, 3, pp.445-453.

Schoenwald, S.K. (1998) *Multisystemic Therapy Consultation Manual*. Charleston, South Carolina: MST Institute.

Schoenwald, S.K. (2008) Toward Evidence-Based Transport of Evidence-Based Treatments: MST as an Example. *Journal of Child and Adolescent Substance Abuse*, *17*, 3, pp.69-91.

Schoenwald, S.K., Chapman, J.E., Sheidow, A.J. and Carter, R.E. (2009) Long-Term Youth Criminal Outcomes in MST Transport: The Impact of Therapist Adherence and Organizational Climate and Structure. *Journal of Clinical Child and Adolescent Psychology, 38*, 1, pp.91-105.

Schoenwald, S.K., Henggeler, S.W. and Edwards, D.L. (1998) *MST Supervisor Adherence Measure* [Unpublished instrument]. Charleston, South Carolina: MST Institute.

Schoenwald, S.K. and Hoagwood, K. (2001) Effectiveness, Transportability, and

Dissemination of Interventions: What Matters When? *Psychiatric Services, 52*, 9, pp.1190-1197.

Schoenwald, S.K., Sheidow, A.J. and Letourneau, E.J. (2004) Toward Effective Quality Assurance in Evidence-Based Practice: Links Between Expert Consultation, Therapist Fidelity, and Child Outcomes. *Journal of Clinical Child and Adolescent Psychology, 33*, 1, pp.94-104.

Scott, S., Knapp, M., Henderson, J. and Maughan, B. (2001) Financial cost of social exclusion: follow up study of antisocial children into adulthood. *British Medical Journal, 323*, pp.191-194.

Scott, S., Sylva, K., Doolan, M., Price, J., Jacobs, B., Crook, C. and Landau, S. (2010) Randomised controlled trial of parent groups for child antisocial behaviour targeting multiple risk factors: the SPOKES project. *Journal of Child Psychology and Psychiatry, 51*, 1, pp.48-57.

Substance Abuse and Mental Health Services Administration (2003) *Using General Organizational Index for Evidence-Based Practices*. Rockville, Maryland: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services.

Substance Abuse and Mental Health Services Administration (2009) *Family Psychoeducation: Evaluating Your Program.* Rockville, Maryland: Center for

Mental Health Services, Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services.

Substance Abuse and Mental Health Services Administration. (2009) *Illness Management and Recovery: Evaluating Your Program.* Rockville, Maryland: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services.

Substance Abuse and Mental Health Services Administration (2009) *Supported Employment: Evaluating Your Program.* Rockville, Maryland: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services.

Sundell, K., Hanson, K., Löfholm, C.A., Olsson, T., Gustle, L.-H. and 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*, 3, pp.550–560.

Suter, J., Burchard, J.D., Force, M., Bruns, E.J. and Mehrtens, K. (2002) *User's manual to the Wraparound Fidelity Index 3.0.* [online]. Available at: http://depts.washington.edu/wrapeval/docs/WFI_Manual.pdf . Accessed 29 March 2012.

Teague, G.B., Bond, G.R. and Drake, R.E. (1998) Program fidelity in assertive community treatment: development and use of a measure. *American Journal of Orthopsychiatry, 68*, pp.216-232.

Turner, K.M.T. and Sanders, M.R. (2006) Dissemination of evidence-based parenting and family support strategies: Learning from the Triple P-Positive Parenting Program system approach. *Aggression and Violent Behavior, 11*, pp.176-193.

Waltz, J., Addis, M.E., Koerner, K. and Jacobson, N.S. (1993) Testing the Integrity of a Psychotherapy Protocol: Assessment of Adherence and Competence. *Journal of Consulting and Clinical Psychology, 61*, 4, pp.620-630.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., Blachman, M., Dunville, R. and Saul, J. (2008) Bridging the Gap Between Prevention Research and Practice: The Interactive Systems Framework for Dissemination and Implementation. *American Journal of Community Psychology, 41*, pp.171-181.

Williams, N., Oberst, J., Campbell, D. and Lancaster, L. (2011) The Children's Psychosocial Rehabilitation Treatment Adherence Measure: Development and Initial Validation. *Community Mental Health Journal, 47*, 3, pp.278-285.

Wilson, D.C. and Crisanti, A.S. (2009) Psychometric Properties of the Dual-Disorder Treatment Fidelity Scale: Inter-Rater Reliability and Concurrent Validity. *Community Mental Health Journal, 45*, pp.171–178.

World Health Organization (2007) *International Classification of Diseases: Chapter V; Mental and behavioural disorders (F00-F99); Behavioural and emotional disorders with onset usually occurring in childhood and adolescence (F90-F98), F91 Conduct Disorders* [online]. Available at: http://apps.who.int/classifications/apps/icd/icd10online/?gf90.htm+f91 . Accessed 18 September 2010.

Zazzali, J., Sherbourne, S., Hoagwood, K.E., Greene, D., Bigley, M.F. and Sexton, T.L. (2008) The Adoption and Implementation of an Evidence Based Practice in Child and Family Mental Health Services Organizations: A Pilot Study of Functional Family Therapy in New York State. *Administration and Policy in Mental Health, 35*, pp.38-49.

Part 2: Empirical Paper

Complex interventions for children and young people with conduct problems: developing a measure to explore service delivery and characterise interventions

Abstract

Aims This paper reports on the development of the Children and Young People – Resources, Evaluation and Systems Schedule (CYPRESS), a measure aimed at characterising the service delivery frameworks of services offering complex interventions to children and young people with conduct and associated problems as part of the Systemic Therapy for At Risk Teens (START) trial comparing multisystemic therapy (MST) with management as usual (MAU).

Method A review of the literature was conducted which informed the development of CYPRESS, specifically as regards the structure and content of a scale of this nature. CYPRESS was further developed on the basis of expert review and piloting. Finally, it was administered to 16 teams taking part in the START trial.

Results Median scores on the CYPRESS measure across the three subsections and total score were significantly higher in the MST teams than MAU, and the range of scores in MST was significantly narrower. A case comparison explored a number of ways in which the highest and lowest scoring MAU teams were different from one another in terms of the service delivery framework in which they work.

Conclusions CYPRESS is a viable means of distinguishing between services providing complex interventions for children and young people. It offers the chance to characterise services on the basis of the service delivery frameworks which support them rather than specific features of interventions, and as such can be used

across a range of different services. Further development and testing of CYPRESS is recommended.

Background

This paper reports on the development, implementation and findings, in its initial application, of the Children and Young People – Resources, Evaluation and Systems Schedule (CYPRESS; see Appendix F), a measure aimed at assessing fidelity in and characterising services offering complex interventions to young people with conduct disorder and associated problems. The use of CYPRESS to characterise both the multisystemic therapy (MST) and management as usual (MAU) treatment conditions constitutes part of the Systemic Therapy for At Risk Teens (START) trial, a large multi-centre randomised controlled trial (RCT) examining the effectiveness of MST, well-established in the United States where it originated, in the UK context. Other trials outside the US have shown variable results, with studies in Canada (Leschied and Cunningham, 2002) and Sweden (Sundell et al., 2008), for example, finding no difference in outcome between MST and MAU. The value of examining MST in the UK context is apparent.

The START trial is a multi-centre endeavour across nine sites in England, in which, to date, 655 young people have been randomised to either MST or MAU. It represents a collaboration between University College London, the Institute of Psychiatry at King's College London, the University of Cambridge and the University of Leeds, and is funded by the Department of Health and the Department for Education. Its aim is to effectively compare MST with existing services available to young people considered to be at high risk of requiring out-of-home care such as fostering, social care placement or, in the case of young people engaging in offending behaviour, custody, and the families of those young people.

Introduction

Recent years have seen huge growth in the value placed on evidence-based practice (EBP), the 'conscientious, explicit and judicious' use of current best evidence to guide clinical decision-making in health and social care (Sackett, Rosenberg, Gray, Haynes and Richardson, 1996, p.71). Of particular interest here is the increasing centrality of EBP in interventions for internalising and externalising disorders in children and young people. MST is an obvious example (Curtis, Ronan and Borduin, 2004); also relevant are parent management training (for a review of the evidence see National Institute for Health and Clinical Excellence, 2006) and cognitive behaviour therapy (CBT) for anxiety disorders in children and young people (In-Albon and Schneider, 2007; James, Soler and Weatherall, 2005). It should be borne in mind that the importance of EBP is bound up with the ever growing focus on outcomes in the NHS and their relationship with funding (see, for example, *Equity and excellence: Liberating the NHS*; Department of Health, 2010). It is therefore essential that the research and clinical communities have robust means of measuring and communicating what interventions look like in both the specific treatment and usual service conditions of research evaluations, and how their components might relate to outcomes.

As the emphasis on evidence-based practice in delivering behavioural, mental health and other interventions continues to grow, the means by which practice is implemented and monitored becomes increasingly important. The expanding fields of diffusion of innovation (see e.g. Greenhalgh et al., 2004; Ferlie and Shortell, 2001; Schoenwald, 2008) and implementation science (see e.g. Damschroder et al., 2009) aim to address the questions around how to effectively disseminate and sustain health service innovations (Greenhalgh et al., 2004). Further, there is an understanding that

it is essential to have means of knowing whether interventions are being accurately replicated when transported from the experimental conditions under which they were developed, and which factors are important in ensuring this. Adherence or fidelity scales – broadly, measures which assess the degree to which specific features of a treatment manual are adhered to, or the extent to which overarching principles of an intervention are demonstrated – are one way in which researchers and practitioners have attempted to introduce standardisation into practice and examine what factors are associated with effective implementation when interventions are transported to different clinical environments.

This is seen as a fundamental tenet of high quality practice and central to ensuring that interventions are delivered as intended, and therefore true representations of the evidence base from which they claim to spring. MST is a 'gold standard' example in terms of rigorous measurement of adherence as it appears in practice, employing as it does the therapist adherence measure (TAM; Henggeler and Borduin, 1992), supervisor adherence measure (SAM; Schoenwald, Henggeler and Edwards, 1998) and consultant adherence measure (CAM; Schoenwald, 1998). Further, the developers of MST have addressed transportability in research (Schoenwald, 2008); cultivating adherence among therapists, supervisors and consultants is one important element, along with organisational factors such as staff participation in decision-making and a clear hierarchy of authority.

While measurement of treatment adherence in terms of what clinicians do in practice is increasingly expected in the delivery of EBP, other features of a given intervention, such as those which support implementation and service delivery, are, by and large, less clearly examined. Service delivery framework elements, such as supervision procedures, team communication processes, outcome monitoring and

referral criteria (features heavily emphasised in MST, as an example), are likely to be of central importance.

Factors which support service delivery

We are interested to see whether these service delivery factors are replicated in therapies other than MST, and to understand the impact of these features of service delivery on outcome. The aim is to establish whether it is possible to measure factors such as this in the same way as specific techniques are measured in the adherence measures which are increasingly used.

The elements of MST which support its practice have been examined widely. Curtis et al. (2004) report a meta-analysis which demonstrates positive outcomes for MST as an intervention to reduce antisocial behaviour in young people. More specifically, they highlight that outcomes are optimal when interventions are applied as intended, indicated by high scores on the TAM. They analysed the magnitude of treatment outcome across seven primary outcome studies and four secondary studies, and while they found favourable outcomes across the board, there were site differences which moderated the effectiveness. Those sites in which supervision was delivered by the MST developers themselves, rather than by supervisors who had been trained by the developers, achieved better outcomes. While other sites were implementing MST in accordance with the manual, it might be argued that there is more room to deviate from the intended intervention, perhaps in terms of aspects of service delivery, in these conditions (those where a developer is not supervising). This serves to highlight the absolute centrality of implementation to effective interventions, and its multifaceted nature.

Effectiveness studies provide opportunities to develop thinking around those factors which are important in successful interventions. While MST has an extensive

body of research which attests to its effectiveness, as noted above there are, nonetheless, exceptions, that is to say large-scale trials which have demonstrated less favourable results. The studies conducted in Canada (Leschied and Cunningham, 2002) and Sweden (Sundell et al., 2008) in recent years both found no advantage of MST over usual services. In the UK, Butler, Baruch, Hickey and Fonagy (2011) have recently explored the value of MST in this particular cultural context, and consider it to be an effective additional intervention for young people exhibiting offending behaviour, recommending further exploration (the START trial being an example of such). The authors draw out the difficulty of identifying the most beneficial aspects of MST, but comment on the ethos and practices of MST as being distinct from that of usual services (youth offending teams in this example). The capacity to measure elements such as ethos, characterised in MST by 24-hour availability of the team and clinicians assuming responsibility for change, would be a valuable addition to existing measurement capabilities.

There are clear implications for the role of implementation in establishing evidence-based practices, and for what we consider to be robust mechanisms for measuring the various factors which relate to successful implementation of an intervention. This includes those factors which might not be essential to maintaining adherence to a particular therapeutic model, but are characteristics of the service context which may be fundamental to its delivery.

It might be argued therefore that the framework which supports the delivery of a given intervention is as important as the model itself and its component parts. MST is an example of how this might be seen in practice: when an organisation purchases a licence from MST Services, Inc., it gains access not only to a model, but to a comprehensive service delivery system that accompanies it. This includes highly

structured and rigorous supervisory and consultation arrangements, a programme of ongoing training, and a framework which articulates the nature of relationships with stakeholders, as well as, crucially, a requirement for practical and financial support structures within the wider organisation which are able to support these practices and ways of working. These are factors which can be seen to support service delivery in complex interventions for young people and their families, but which transcend specific theoretical models.

If we are to evolve evidence-based practice which is truly replicable and which can be seen to genuinely offer the interventions that the relevant evidence base proposes, it seems necessary to have means of measuring the degree to which providers of interventions implement the relevant service delivery frameworks, just as we concern ourselves with the way in which they implement specific therapeutic techniques and approaches to working with clients.

Examples of existing measures of service delivery

In the adult field, the Dartmouth Assertive Community Treatment Scale (DACTS; Teague, Bond and Drake, 1998) exists as a formal fidelity measure of the service delivery aspects of one particular type of intervention, assertive community treatment (ACT). The measure looks at three broad areas: human resources; organisational boundaries; and nature of services. The DACTS specifies that it measures fidelity at the team level, rather than an individual or agency level. It uses a multi-informant approach, drawing on interviews with staff working at different levels in the organisation, and written information. While the DACTS instructions indicate that it can be used with groups other than ACT teams, it is specifically designed for ACT services and the degree to which the developers would recommend use with non-ACT teams is a little unclear.

Contrasting with the above, which is designed for adults with severe mental health problems, the Youth Integrated Community Treatment Fidelity Scale (YICT) (adapted from the Integrated Dual Disorders Treatment Fidelity Scale; Mueser Noordsy, Drake and Fox, 2003) is a measure which is designed for complex interventions with children and young people with comorbid mental health and substance use problems. The YICT addresses some service delivery elements, such as a multi-disciplinary approach, team meetings and community and home-based provision of services, but also largely specifies and measures particular modes of intervention or focuses for treatment, for example, motivational interviewing, anger management and CBT. This allows for specific examination of the given intervention, but is not applicable across different therapeutic contexts.

Developing a measure of service delivery

The intention here was to develop a fidelity measure which allows assessment of the service delivery aspects of complex interventions for children and young people with conduct and related psychosocial problems, and to characterise MST and MAU in the START trial along these dimensions. The measure is intended to have utility across a range of interventions in the wider child field, developed as it has been to capture the service delivery frameworks of both MST and usual services.

A central consideration is which aspects of service delivery should be included in such a measure; if the factors investigated do not reflect that which we know to be important in effective practice then the measure is redundant no matter how well it might be designed or executed.

A report for the Department of Health and Prime Minister's Strategy Unit which reviewed interventions for children at risk of developing antisocial personality disorder (Utting, Monteiro and Ghate, 2007), offers valuable considerations on which

to draw. Utting and his colleagues examined the available evidence for six existing interventions for young people with conduct problems, known to have achieved positive outcomes internationally as demonstrated by the literature, and which are available in the UK (two parenting programmes, *The Incredible Years* and *Triple P*; the *Nurse-Family Partnership* home visiting programme; and three programmes for families and carers of high-need children and adolescents, *MST, multidimensional treatment foster care* (MTFC) and *functional family therapy* (FFT)). One exception to the inclusion criteria (known to have achieved positive outcomes internationally as demonstrated by the literature, and available in the UK) was Functional Family Therapy (FFT), which has not been evaluated with the rigour that the others have, but which has been identified as achieving positive outcomes with the target group, and has some history of use in the UK.

From their examination of the literature, Utting et al. (2007) derive the following key factors in effective implementation, found to be present in all the interventions considered:

- A strong, coherent and clearly articulated theoretical basis
- Professional, qualified and/or trained staff
- High programme fidelity
- Delivery of interventions in the natural environments of children and young people
- Interventions which are tailored to the needs of individual clients
- Partnership with families
- A multi-modal or multi-dimensional approach
- A sustained/intensive approach

In addition to the above factors, it is also important to consider more broadly the elements which practice tells us are important to well functioning services to address the complex needs of children and families, such as continuing professional development (see, for example, the British Psychological Society's *Continuing Professional Development Guidelines*, 2010); quality improvement (Department of Health, 2008); factors highlighted in *Every Child Matters* (Department for Education and Skills, 2003), such as comprehensive assessment procedures, effective risk management, well defined relationships to other services, and regular and effective procedures for communicating with other professionals and people involved in a young person's care; and features emphasised in Roth, Calder and Pilling's competence framework for child and adolescent mental health services (2011) such as high quality supervision and case coordination. Finally, we can also draw on the literature which describes MST (see Henggeler, 1999; Henggeler and Borduin, 1990) and underlines the importance of clarity on practical organisational issues such as the nature of the population served and service capacity; well defined care pathways; and clearly articulated professional roles.

Based on this literature, a measure was designed to assess the key elements or dimensions of service delivery frameworks that have been identified as central to effective implementation of complex interventions for children with conduct problems and their families. By applying this measure, we will ultimately be able to test the hypothesis that there is a relationship between these service delivery factors and outcomes as indicated by performance on measures of recidivism, out-of-home placement, education and others measured in the START trial. Second, by applying this measure to the MAU condition in the trial, it is possible to better characterise MAU, which is largely under described in published RCTs implemented with

children and families. MST is an example of a body of literature where numerous studies report on MAU but do not sufficiently describe it. In Canada and Sweden, where less favourable results were found for MST (Leschied and Cunningham, 2002; Sundell et al., 2008), there is an argument that MAU services in those countries are fundamentally different from MAU in others where MST has been successful. Without accurate characterisation of MAU we cannot address this possibility.

As the multi-site START trial is still ongoing, we are prohibited from reporting outcome data in this paper. Consequently, our focus will be to apply CYPRESS to characterise both MST, and, importantly, MAU services along important service delivery dimensions usually neglected in research of this nature, and to consider whether CYPRESS is able to distinguish between the two conditions.

It might be the case that MAU interventions do not possess the elements considered important from a service delivery perspective – though indeed some might possess them and they might be effective – but without accurate characterisation of them we are essentially comparing one highly specified condition with one about which we know very little. The purpose of this study is to look at the quality of the service frameworks which support both MST and MAU in this large trial, offering as it does a window on an intervention which is both highly structured and relatively recently implemented (in the UK), and the existing alternatives, which are largely unknown.

Logically, adherence measures are derived from the manuals which guide the implementation of a given intervention and are consequently intervention-specific. CYPRESS will allow us to take the concept of fidelity a step further by offering a useful measure which incorporates elements which are not specific to particular therapeutic models but rather which are seen across different interventions serving

the same population. CYPRESS was conceived as a means of examining the service delivery frameworks that surround a given intervention and which might be as important in successful implementation and dissemination as the specific components of individual sessions informed by model-specific practices. It allows interesting comparison of the service delivery frameworks which exist in different types of services, offering interventions to the particular population of children and young people experiencing conduct or related difficulties across settings. Further, the measure offers a means of making this comparison, and characterising these services in a novel way, in a manner which is replicable.

To state the focus of this research in terms of research questions, we have aimed to establish the following:

- What are the service delivery characteristics of complex interventions for children and young people that relate to outcome and can they be reliably measured across MST and MAU?
- What is the nature of the many MAU services which characterise the comparison arm of the START trial?

Method

This piece of work entailed the development, piloting and implementation of a new measure, CYPRESS, aimed at characterising services along service delivery framework lines. A further aim was to use the CYPRESS procedure to gather information which would allow us to characterise MAU services offering interventions to young people with complex presentations.

Ethical approval

Ethical approval for a fidelity measure aimed at characterising the trial interventions was given as a feature of the initial permission for the START trial. This was granted by the Research Ethics Committee of the National Research Ethics Service South East (see Appendix G).

Participants in this research were clinicians being interviewed in a professional capacity about the service within which they work. It was not considered likely that the participation of staff would result in any distress to them, the effects of which it would be incumbent upon us to mitigate. Clients of the services under examination were not involved in the process, nor was case-specific clinical information collected.

Measure development

Specifying the scope. The basis for CYPRESS was conceived in the planning of the START trial (prior to this author's involvement in the trial) as a means of characterising key elements of functioning in MST and MAU in terms of how interventions are implemented. Utting et al.'s (2007) key factors constituted a theoretical grounding for a number of the elements to be incorporated into the measure. A literature review, described in part one of this volume, was conducted which allowed the CYPRESS developers to consider what existing measures or other documents might offer in terms of understanding the measure's appropriate form and content. Established professional practice constructs were also considered, for example, supervision, continuing professional development, and team communication, along with key issues in modern services across health and social care, such as service capacity, assessment procedures and risk management.

Designing the measure. Having specified the scope and the constructs to be included, through review of existing measures, literature on conceptual frameworks and practice in the field, and insights of the review by Utting et al. (2007) on the necessary features of interventions for young people and their families, the developers were in a position to generate a draft measure.

The measure was developed by Stephen Pilling (Clinical Psychologist with expertise in evidence-based practice and experience of measure development), Stephen Butler (Child Clinical Psychologist with expertise in the area of complex interventions for young people with conduct disorders and the author's supervisor) and Cressida Gaffney (Trainee Clinical Psychologist), based on principles of effective interventions and drawing on knowledge of existing measures. Having identified three overarching domains to be addressed in the measure, the developers drafted a list of potential items for each domain; this was subsequently refined to include 20 items considered to reflect the areas which emerged from the review of the literature as important for understanding service delivery frameworks supporting complex interventions.

The three sections of CYPRESS, aimed at capturing key issues pertaining to service delivery level factors that characterise a service, are as follows: service characteristics (scored out of 30), team operation (scored out of 30) and delivery of interventions (scored out of 40) (see Appendix F). Below are the item headings and one sample item for each of the three sections:

Service characteristics. Item headings are as follows: 'shared model of care'; 'population served'; 'care pathway'; 'service capacity'; 'relationship to other services' and 'service/team staffing'.

Sample item: 'Shared model of care'

'Service has a comprehensive and shared view of the model of care provided which is owned by the service.'

Team operation. Item headings are as follows: 'team meetings'; 'supervision'; 'staff training'; 'team communication'; 'client outcome monitoring' and 'quality assurance'.

Sample item: 'Team communication'

'Team has clear policies and procedures for communicating information about clients and decisions made by the team, with team colleagues, to clients and with other agencies.'

Delivery of interventions. Item headings are as follows: 'range of interventions consistent with model'; 'assessment'; 'individualised care'; 'family/carer involvement'; 'assertive engagement'; 'interventions provided in a range of settings'; 'risk and child protection' and 'case management'.

Sample item: 'Assertive engagement'

'The team has an assertive approach to the engagement of clients and families/carers, (e.g., a 'no drop out' policy or a stress on overcoming difficulties in engaging with services).'

The opinion of a further two experienced clinicians was called upon in the development of the measure; they reviewed an early draft, and their considerations were incorporated into the design of CYPRESS. Time constraints meant that the planned modified Delphi approach was not possible in its entirety; this will be considered in our discussion.

Piloting. The schedule was piloted through interviews with team members and team leaders of two non-START trial services offering complex interventions for young people (one MST team and one team representing MAU, a multi-agency team

offering intensive support to young people with high levels of complex needs and challenging behaviour, usually where there is a risk or history of family or foster placement breakdown). In both cases the population served was considered similar to those being seen by the teams to be interviewed for the trial. These teams were identified and approached by one of the measure developers (SB) to whom the service managers of the teams were known. All three developers were present for the pilot interviews. Notes on issues related to the content of the measure and process of its administration were made during the pilot phase. A meeting was held in which the developers refined the measure on the basis of the pilot, addressing each measure item in turn, followed by general theoretical and pragmatic considerations raised in the piloting process.

Scoring. CYPRESS was designed to be accessible and simple to use. The scale was designed to be scored out of 100, with 20 questions to be scored on a five point scale, from one to five.

The scale is designed in such a way that the possible scores are given alongside a brief description of the features that characterise that particular score, so raters have the specific features of a given rating available at the time of scoring. Further, the CYPRESS manual (see Appendix H) was developed to guide raters where further clarification was needed as regards the scoring scheme. This was developed by one of the three scale developers (SP) and subsequently revised by all three developers to correspond with the scale as it was improved.

The presence of two raters was key, and agreement between them on a final score considered important for reliability. Rating by two people was warranted due to the complexity of the data, and to reduce the chance of bias in scoring, thus increasing the robustness of the scores allocated to teams.

Reliability rating process. In order to assess the reliability of the schedule, a second pair of independent raters, an experienced Child Clinical Psychologist and a more newly qualified Clinical Psychologist not involved in the START trial or development of CYPRESS, will conduct reliability testing. These raters will listen to audio recordings of a sample of the interviews and rate each team on the basis of the interviews, using CYPRESS. Their ratings will then be compared with those made by the primary raters and statistical comparison conducted. Reliability rating is a fundamental element of the process in measure development and as such will be carried out on CYPRESS during the period of the START trial. It was unfortunately not implemented in time to be reported in this paper.

Psychometric properties. The requirements of the wider START trial meant that it was not possible to conduct psychometric analysis of CYPRESS within the timeframe necessary for the completion of this paper. However, a clear next step, should the establishment of CYPRESS as a viable tool be possible, would be comprehensive analysis of its psychometric properties.

Participants

Staff working in eight MST and eight MAU services (one MST and one MAU team in each of eight locations across England) with children and young people with complex needs constituted the participants in this piece of research (one of the nine START trial sites had to be excluded due to contact not being established with the relevant MAU team despite repeated efforts). The results section will go on to elucidate the nature of the services from which the participants were drawn as this was in fact a large element of what we sought to investigate. The respondents were staff working either in a ‘team member’ or ‘therapist’ capacity, or in a ‘team manager’, ‘team leader’ or ‘supervisor’ capacity in these services; for brevity

participants will henceforth be referred to as ‘therapist’ or ‘supervisor’. Decisions as to which specific therapists (and supervisors where there was more than one) participated were made by the teams themselves and appeared to be largely based on availability. The services can be broadly described as providing interventions to young people with complex needs, primarily but not exclusively around behavioural issues, with complex family difficulties often a central concern. As we will see, these services differed considerably in a number of ways, but the uniting feature was that they offered some sort of intervention to children and young people with complex needs, often around offending behaviour or high levels of challenging behaviour in the family home. Services came to be in the trial by virtue of being either an MST team, or a commonly utilised MAU service which participants in the START trial might be offered if not randomised to MST.

The MST teams interviewed were already engaged in the START trial, indeed, these teams were established as part of the trial; they had been providing MST clinical services to families for between approximately 18 and 24 months prior to implementation of the RCT. The MAU teams were specified by the relevant MST teams, who were asked to identify the most commonly used MAU in their geographical area, that is to say the service most frequently offered to the young people not randomised to MST. Where there was uncertainty on the part of the MST teams as to what service constituted the most commonly used MAU service in the area, the START trial coordinator was asked to establish from existing MAU data which service would be most appropriate.

Procedures

Setting up the interviews. As outlined above, the participants were members of teams (either MST or MAU) involved in the START trial. As a first step in

disseminating information about the rollout of the measure, a presentation was given at a meeting of representatives from all the MST teams involved in the trial, by one of the START principal investigators with a specific role in developing the measure (SP). This outlined the rationale for the development and implementation of the measure, along with the associated proposed timescale. Absent from this meeting were representatives of MAU, as it was open only to MST teams and the START research team.

One of the CYPRESS developers (CG) liaised with all MST and MAU teams to advise on the requirements of the CYPRESS process, arrange site visits and request completion of pre-interview information (see below). Between November 2011 and March 2012, the MST teams were visited and the interviews conducted. Concurrently, the MST teams were asked to identify their local MAU services. Once this information was shared with the researchers, the MAU teams were contacted, by email in the first instance and by telephone following this where required, and interviews arranged and conducted (December 2011-April 2012).

Prior to interviews, the CYPRESS pre-interview information form (see Appendix I), requesting brief factual information about the team's organisational context, staffing, and supervision and meeting arrangements, and copies of documentation such as operational policy, where available, was sent to teams by email and returned to the author. Having received the pre-interview information, the raters (SB and CG in the case of all but two interviews, see below) reviewed it and identified any areas where minimal questioning would be required as the area was felt to be covered fairly comprehensively in the documentation, and the converse, those areas where particular attention was needed due to the absence of information or the need for clarification. In the case of three MAU teams, information was

received after the CYPRESS interviews had taken place, meaning that the pre-interview information review and planning was not possible in these cases. The CYPRESS manual was available to raters at all stages to aid with planning and conducting the interviews, and scoring the measure.

Conducting the interviews. For the administration of the interviews, a site visit was conducted by the raters (two of the three CYPRESS developers). During this visit, two interviews were conducted, one with the supervisor and a separate interview with two therapists. In the case of three teams (one MST and two MAU) only one therapist was available for interview. No stipulation was made as regards whether supervisor or therapists were interviewed first, meaning that interviews were conducted in the order that was most convenient to the teams. Prior to the interview participants were provided with the CYPRESS participant information sheet (see Appendix J), and given a verbal explanation of the purpose of the measure and its role within the START trial, along with the opportunity to ask any questions they might have.

All interviews were audio recorded for reliability rating at a later stage (audio recordings were stored on a password protected computer and will be deleted after reliability scoring). 30 of the 32 interviews were led by one of the two experienced clinical psychologists (28 by SB and two by SP), and the remaining two were led by the author (CG) (the non-leading rater supplemented the first rater's questions as required according to their judgement). During interviewing both raters gave independent provisional scores and the non-leading rater made notes on the content of discussion.

Agreeing final ratings. After completion of both interviews for a given site, raters reviewed together their provisional scores for therapist and supervisor

interviews and agreed on a final team score for each item. Where there was disagreement the written notes were reviewed and each rater offered their reasoning for a given score, in order to reach agreement. The CYPRESS manual stipulates that if this does not lead to resolution, the next step in the process is to seek clarification or more information from the team in question. As a final step, raters have the option of taking the query to a third party. In the event, no disagreements were encountered which required further information from the team, or advice from a third party. Indeed, there was often close agreement between raters based on their independent ratings.

Feedback to teams. Brief feedback will be provided to participating teams; this will be a general summary of our findings on what usual services look like across the sites investigated, and how the measure has informed our understanding of the service delivery frameworks which are in place across the teams participating in the trial.

Characterising MAU

CYPRESS was used to characterise the teams interviewed. Analysis of the scores allowed clustering of services on the basis of performance on the measure. A case comparison of the two lowest scoring and two highest scoring MAU services was conducted. Total score was used as the means of distinguishing different categories of service. The case comparison was supplemented with information gathered during the interview process and from the pre-interview information form.

Results

Characteristics of the START sample

While the young people participating in the START trial did not constitute the participants in this specific piece of work, they nevertheless constitute the clinical population served by the teams under examination and as such the characteristics of the sample should be noted here.

The children and young people in the START trial (n = 655) all lived in England. Participants' ages ranged from 11 to 17 years, with a median age of 13.8 years. 63% were male (n = 415). Where ethnicity data was recorded at the time of writing (n = 456), the breakdown was as follows: 77.6% white British or white other (n = 354), 11.6% black British or black other (n = 53), 7% mixed (n = 32), 3.3% British Asian or Asian other (n = 15), 0.4% other (n = 2). In terms of the stage of onset of conduct problems, 57% (n = 373) were classed as late onset (at the age of 11 years or older), with the remaining 43% (n = 282) classed as early onset. 39% (n = 255) of the sample had an officially recorded criminal offence. 17% of the sample (n = 110) were classified as having special educational needs.

Analysis of CYPRESS scores

Group differences between MST and MAU on the three subsection scores and total score were analysed. Given that each geographical location has two sites serving the same population (one MST and one MAU), the data have been treated as matched pairs. A non-parametric paired samples Wilcoxon signed-rank test was used to compare means; this test was selected as the data are scores, and derived from a small sample. Given the small sample size, this was complemented with a Monte Carlo simulation to ensure the robustness of the findings.

Table 1 illustrates the median scores and range of scores on CYPRESS for the MST and MAU teams. The MST teams consistently showed higher median scores and narrower ranges of scores across the three subsections and total score.

Table 1 CYPRESS median scores and ranges

Subsection		MST	MAU
Service characteristics	Median	27	21.5
	Range	4	7
Team operation	Median	25	20.5
	Range	5	6
Delivery of interventions	Median	35.9	29
	Range	5	11
Overall score	Median	87.5	70.5
	Range	13	24

There was a significant effect of group, $z = -2.524$, $p = 0.01$, $r = 0.89$, with higher scores in the MST teams than the MAU teams. A two tailed Monte Carlo test with a confidence interval of 99% estimated a highly significant p value of 0.006.

Characterisation of MAU

The teams constituting MAU represented the following categories:

- Child and Adolescent Mental Health Service
- Youth Offending Team
- Family support charity project
- Specialist local authority child safeguarding service
- Specialist local authority family intervention team

In order to offer a characterisation of MAU, a team comparison will be described, which looks at the two lowest and two highest scoring MAU services, with an exploratory, broadly phenomenological stance, drawing in part on thinking from research on organisations in business studies (see Ghauri, 2004). This comparison employs review of both CYPRESS scores and the interview notes compiled by the second rater. Table 2 shows the scores of the two lowest and two highest scoring MAU teams.

Table 2 The lowest and highest scoring MAU teams

	Lowest scoring		**Highest scoring**	
Subsection	**Team A**	**Team B**	**Team C**	**Team D**
Service characteristics	18	18	23	25
Team operation	20	20	25	26
Delivery of interventions	28	27	33	38
Total	**66**	**65**	**81**	**89**

The pattern seen in total scores is reflected throughout the subsections, with teams A and B scoring lower on all sections than teams C and D.

While the services constituting MAU were various, there are themes which emerge from consideration of the interviews with the teams which demonstrate common features of the teams which scored less well and those which scored more highly, extending the understanding gained from the CYPRESS scores alone. The highest performing MAU teams were different from the lowest in a number of ways, which will be considered below.

Service characteristics. Teams C and D were both local authority services, while teams A and B were services which were jointly commissioned by NHS or charity services, and the local authority. The dual or shared nature of these services, and the tensions arising from this, were apparent in the interviewees' descriptions throughout the CYPRESS procedure, in relation to, for example, child protection policy, team communication, and training.

Teams C and D described clearly articulated philosophies of care which were reported to be shared by the team, while teams A and B reported different approaches to care depending on the client in question. The latter both said that to have a single view of the model of care would be inappropriate given the diverse nature of problems served by the teams. This is a fundamental point linked to the referral criteria – while the lowest scoring teams served very broad clinical populations, the highest scoring teams served very specific groups. Referral reasons in the highest functioning teams were likely to represent problems of a nature both more serious and more complex, with higher numbers of child protection cases and serious difficulties disrupting the family home.

Teams A and B offered less clarity around care pathways in their responses, and this again might be seen to be related in part to the diverse populations served, in that different care pathways might be appropriate depending on the nature of the referral. Both these teams commented that there was a lack of clarity around when and how to end an intervention. Team C had a clear and robust care pathway which specified staff actions at all stages throughout it. Team D offered an unusual perspective, commenting that they did not have a clear pathway, but that this was by design, with workers applying flexibly means of working with a particular family. This service was in a sense defined by its flexibility, and yet operated within a clear

set of boundaries and with a high level of knowledge and skill as regards when and how to employ particular approaches or interventions. Indeed, this was the highest scoring MAU by some points.

Clarity around questions on maximum caseload numbers, that is both whether such a maximum existed and what it was, was evident in teams C and D. Moreover, there was agreement between supervisors and therapists on both of these questions. In team B, there was inconsistency as to whether caseload maximums were appropriate or possible. These issues could not clearly be linked to or explained by the actual caseloads; teams A, C and D had very similar individual caseloads (ranging from 8-15); with team B having notably higher individual caseloads of approximately 35.

Clearly defined roles and explicit supervisory hierarchies (distinct from the question specifically on supervision, addressed elsewhere) were most apparent in teams C and D. Contrasting with teams A and B, the higher performing teams comprehensively articulated the nature of the different roles in the team, the ways in which they interact with one another, the formal hierarchies which govern lines of responsibility and decision-making in the team, and detailed descriptions of when and how particular responsibilities are enacted. There was variation in terms of the requirement for professional qualifications across the teams, but what was apparent in the higher performing teams was that where less highly skilled workers were employed, their roles were clearly delineated and their supervision by professionally qualified staff considered integral to their role. The presence of an adequate number of professionally qualified staff might be said to be a necessary but not sufficient condition for a team to be high functioning; team B had a highly skilled workforce as regards professional qualifications, but the utilisation of the different skills did not

appear to be optimal, and role clarity was not apparent, as it was in the highest performing teams.

Team operation. Team meetings and supervision scores were consistently high; all four teams scored 4 or 5 on the items measuring these two constructs. Questions as regards the sensitivity of item 8 (supervision), in particular, will be addressed in the discussion, as teams across the sample scored highly with little differentiation, yet there were examples of less and more effective supervisory practices which were reflected in respondents' answers but which could not be accurately be reflected in scoring because of the scope of the measure.

Attitudes to training differed in two key ways in the two halves of this MAU sample. Firstly, in teams C and D there were clear criteria around qualifications required to work in the team, rather than any of several qualifications being sufficient to work in the team, these teams also made explicit the clear links between the qualifications required and the job role. The second issue around training was that the higher performing teams described ongoing training as more central to their functioning, with evident support for training from management and the organisation, including for internal training and external courses and qualifications, while in the lower performing teams the question of resources, that is to say the degree to which the organisation supports training financially, loomed larger. Having said that, it was apparent in all the teams that training is somewhat more variably offered in MAU than in MST teams, where there are very clear expectations for training at induction and regular booster sessions for all. Although a range of professional backgrounds qualify someone to work as an MST therapist, there is a compulsory initial training which must be undertaken before they can commence

work as an MST therapist. In none of the MAU services was the same uniformity of training opportunities seen.

Delivery of interventions. Assessment was an area in which the difference between the higher and lower performing teams was particularly apparent. Teams C and D reported comprehensive, structured assessments which routinely used multiple sources of information. Assessment practices in the other teams showed more variability, and were characterised by a less systemic approach. Involvement of all appropriate stakeholders wherever possible was routine in teams C and D.

What distinguished the lower teams from the higher as regards interventions was not simply whether there was a range of interventions offered by the team, although this was necessary, but whether those interventions were available in a timescale that would be appropriate and therefore useful. Further, there was the issue of which professional a child was allocated to, as not all professionals were able to offer the same service, even when functioning in a generic professional role. This is distinct from those instances where children receive different professional input by design, in order to access a particular intervention, for example, family therapy sessions.

Risk management and child protection was the other area, along with meetings and supervision, where all four teams scored 4 or 5. Risk management and child protection policies are now absolutely standard across health and social care and this was reflected in this question. However, those teams which scored less highly noted some of the challenges arising from its implementation where there are slightly different policies in place across two organisations, and a team is ultimately accountable to both.

The approach to family involvement differed across teams. Teams A and B reported more variability, while teams C and D expressed the centrality of parent or carer involvement. That is not to say that families are not involved heavily in the work of most teams, but the degree to which a focus on family involvement was central differed in the two halves of the sample.

An assertive approach to engagement was observed in teams A, C and D, with only team B scoring less highly on this item. This might suggest that this is a feature of practice which cannot distinguish between services. By and large, assertive engagement was considered an important part of the work with children or young people and their families by all teams interviewed, with the emphasis on it increasing with increasing levels of severity and risk. On the basis of this small sample, it seemed that the more generic a service, the less likely assertive engagement was to be reported.

General observations. Teams A and B were more likely than teams C and D to report that systems or procedures were under development or evolving; there was a sense in which managers were aware in the former services of where they wanted the service to be as regards, for example, use of routine outcome monitoring or having sufficient resource to deliver the desired service, but felt they were not yet at that point. This was not necessarily related to the time for which the team had been in operation – team B had been in existence in their current form for some years longer than both team C and team D, while team A had been in existence for longer than team D.

A particularly striking feature of the teams was the degree to which there was agreement between therapists and supervisors. In the highest scoring teams, there was evident shared knowledge and ways of thinking, while in the lowest scoring

teams there were outright inconsistencies in factual information and an absence of a shared way of talking about the service, both in practical and philosophical terms. The ethos of the higher performing teams was apparent through shared language.

A model of shared knowledge of the cases in the team, albeit at differing levels of detail, was employed in teams C and D, with teams A and B operating a system where a family is known to the assigned worker and supervisor only. None of the teams operated this shared case knowledge model to the same extent as it is used in MST, but it was nonetheless in place in the higher functioning teams.

Teams C and D reported regularly coming together as a team, both in an ad hoc fashion as required and in formal meetings, to share clinical and other professional information. While regularity of team meetings is covered in CYPRESS (item 7), the qualitative difference in how the lowest and highest functioning teams valued and practised coming together as a team, varied in a manner which was not captured by the measure.

The sense of the lowest scoring MAU services was of more disparate, less cohesive teams, with a broader, less well delineated remit or statement of population served. These features did not seem here to be related to the size of the team or indeed of the wider organisation.

The lower performing teams were not equivalent in terms of qualification. However, what was apparent was that whether because qualifications were at a lower level, or because staff were not being utilised in a way that made the most effective use of specific qualifications, there was less clarity around the specific skills expected to do the job in question. There was in both the lower scoring teams a sense of a lack of association between the qualifications or skills of the staff, and the tasks

required of them. To reiterate, this was not only to do with level of qualification but with how the staff members' skills were utilised within the team.

Relationship to outcome

Group differences as regards outcomes cannot be reported here because all outcome data related to the START trial is embargoed until the trial outcomes are due to be reported in full in January 2014.

Due to the minimisation procedure utilised in the trial, it is considered unlikely that there will be differences between MST and MAU in terms of the characteristics of the children and young people in the two groups; participants were matched on age, gender and severity of offending. This means that at such a time as we are able to compare CYPRESS scores with outcomes, it will be arguable that any significant differences in outcome between the two groups are more likely to be explicable in terms of which service was received, MST or MAU, rather than because of differences in the nature of the population.

Discussion

This study brings a novel approach to understanding the features of service delivery which characterise complex interventions for children and young people, considering what these might be and how they can be measured, as well as what the application of CYPRESS tells us in this instance about MAU in the START trial. The results will be considered in terms of what they tell us about measuring these constructs and about MAU. Methodological limitations, implications for future use of the measure, and necessary developments will be considered.

Statistical analyses indicate significant differences between the MST and MAU teams, demonstrating that the service delivery aspects measured by CYPRESS

are more highly developed in MST than in MAU. Across all domains – service characteristics, team operation and delivery of interventions – MST teams scored more highly on average than the MAU services. Further, the MST teams showed consistently higher scores, with less variability across the sites (13 points for overall score compared to 24 in the MAU teams). The higher scoring teams can be said to demonstrate higher levels of fidelity to the service delivery principles underpinning complex interventions for children and young people, as defined by CYPRESS.

These findings might be explained in terms of what is known about MST; as discussed earlier, inherent in the MST model are quality assurance provisions that govern implementation and the approach to the work of the team. Specifically, MST includes factors such as strict procedures for collection of adherence scores (Henggeler and Borduin, 1992; Schoenwald, 1998; Schoenwald, Henggeler and Edwards, 1998), the consultant role, which offers off-site expertise, a model of all staff knowing each other's cases, and strict and clear lines of accountability (see Henggeler, 1999). The organisational features of the MST enterprise might be said to be more akin to a business than the sorts of organisations usually seen in health and social care contexts in the UK. The uniformity apparent in the way MST is set up and delivered in practice is reflected in the relative homogeneity in service delivery characteristics as shown by in the narrower range of the MST scores as compared with MAU.

As we have seen, usual services represent a more disparate group than the MST teams and, as noted, scored consistently lower on measures of service characteristics, team operation and delivery of interventions. The case analysis of the highest and lowest scoring MAU teams allowed for more consideration of the

differences between these two ends of the spectrum and how they might be understood.

Understanding MAU

Characterisation of MAU is an interesting part of this endeavour, offering a perspective on the relatively unexplored usual services, seen here in the context of comparison with MST. Usual services are little described in terms of therapeutic techniques and approaches (Hoagwood and Kolko, 2009; Garland, Brookman-Frazee et al., 2010; Garland, Hurlburt, Brookman-Frazee, Taylor and Accurso, 2010), less still in terms of the service delivery frameworks which underpin them. Some key observations are considered further here.

Organisational boundaries. The distinction between the highest and lowest scoring MAU services as regards whether they were jointly commissioned (for example by the NHS and the local authority) or otherwise, raises important questions for our understanding of multi-agency commissioning of services. The role and specific impact of joint arrangements on team functioning is not clear cut; the challenges of effective joint commissioning or partnership working in health and social care are acknowledged (see e.g. Horwath and Morrison 2007; Hudson, Hardy, Henwood and Wistow, 1999; Leadbetter, 2008). Intuitively, being answerable to two organisations seems likely to impact on team identity and the clarity around service delivery frameworks. This issue is perhaps never more salient, with the impending changes in commissioning which will come into play as a result of the Health and Social Care Act (2012), and it is helpful to be mindful of this. However, we cannot be clear in this instance that joint commissioning arrangements in particular impacted on team functioning, given that many services in health and social care cross organisational boundaries. Indeed, many MST teams, shown here to perform more

highly on the service delivery elements measured, are examples of services which have a background of joint commissioning.

Role clarity and using staff skills. The differences seen in the MAU services in terms of clarity around roles and optimal utilisation of professional skills were informative. The use of a range of staff from different professional backgrounds for generic roles was a feature of the lower performing MAU teams more than the higher, and this is something which anecdotal evidence suggests is common. While there are some skills which are reasonably expected of any professional working in, for example, a child and adolescent mental health service (CAMHS), there are nonetheless likely to be significant differences in the nature of assessment or intervention delivered by, for example, a psychologist, a psychiatrist or a family therapist. Indeed, an inventory of different skills is one of the perceived strengths of multi-disciplinary teams. The tension between using resources creatively and preserving professional identities (Rose, 2011) was apparent here. In contrast, those MAU teams which scored more highly on CYPRESS demonstrated that they were able to make efficient use of the skill mix of the team; moreover, they were set up in such a way that ensuring appropriate resource to carry out specific interventions, in accordance with the relevant model of care, was integral.

Related to this is the value of clear hierarchies of responsibility, at both a day-to-day level and a more strategic level, which was particularly noticeable in comparing the highest and lowest performing MAU teams. As outlined in the reporting of the results, differences were apparent both in terms of explicit responses to questioning about staff roles and in a more abstracted sense, which emerged from the whole interview process. The higher scoring teams' heightened clarity around responsibility, and lines of management and supervision, came through in the

discussion of many of the CYPRESS items. Staff in the higher performing teams were consistently able to identify particular responsibilities and areas of expertise within the team, and these were reported in the same way by therapists and supervisors. These teams answered with more specific details when asked questions about, for example, supervision, risk management and care pathways, and were able to talk about not only the details of a given topic, but, of relevance here, about who would be responsible for different parts of a given process and how these responsibilities would be enacted.

Population served. The breadth or otherwise of the population served was a defining feature of the MAU teams, and as discussed the narrower the referral criteria the more highly specified the service delivery frameworks were. Heterogeneity of population appeared to dilute the specificity of services and seemed to be reflected in the care pathways and in the interventions offered. Those teams which operated with clear and specific referral criteria were able to speak more cogently about the precise nature of the work they offered to young people and their families, the way in which it was organised, and the principles by which it was underpinned. Associated with this is the question of whether services are preventative or curative. While complex interventions might be appropriate in both contexts, variations in approach are likely to be apparent between the two. Service delivery elements are likely to vary depending on which of these roles services consider themselves to be in. The degree to which, for example, a range of interventions is considered essential might vary; services working with people after problems have reached a 'crisis point' perhaps might consider a wide range of interventions less essential than offering a small number of targeted interventions known to be effective with a given population. This is one hypothetical example, but

it exemplifies the subtleties of service provision which should be borne in mind in understanding which features of service delivery might be most salient in different circumstances.

Training. Limited financial or human resources are clearly important in terms of an organisation's attitude to training and the way in which that attitude manifests. A recent King's Fund report (Imison, Buchan and Xavier, 2009) put the annual cost of training in the NHS, for example, at £4bn; this should be understood in the current context of the NHS being asked to find efficiency savings of £20bn by 2014 (Department of Health, 2010). However, it seemed from exploration of the highest and lowest scoring MAU services that the opportunities for training were related to more than resource implications alone. There appeared to be an attitudinal difference between the teams (which was likely to be reflective of organisational attitudes in part; however we cannot dismiss the idea that particular personnel, in terms of other CYPRESS items as well as this, might be an important mediating factor). Without economic comparison of the organisations in question it would be impossible to test this theory, but the positive attitude of the higher performing teams towards training remains an interesting, if not entirely tangible, feature of these teams, which would bear further research.

Assessment. A focus on comprehensive assessment is increasingly evident in children's services, however, the differences seen in terms of the higher and lower performing MAU teams would suggest that this principle is differentially applied. While there are likely to be basic elements that any professional would consider integral to the assessment process, staff from different professional and organisational backgrounds are likely to consider different elements important – a social worker's assessment might look very different from a psychiatrist's. Further,

the frameworks within which assessments are conducted varied widely between teams, with some emphasising strict pro forma, or consideration of a range of sources of information, and others not.

Shared knowledge and cohesion. The concordance of views between therapists and supervisors in the highest performing teams might be said to reflect shared knowledge and expertise, as well as a more personal element of positive interaction between team members. This did not appear to be simply a function of the interpersonal relationships of individuals, though clearly this cannot be entirely disentangled from the question of team cohesiveness and concord, but rather a more fundamental feature of the team in terms of its underlying structure, philosophy and practices. It might be said that task cohesion (the degree to which staff are united and committed to achieving the work task), which has been demonstrated to be positively related to team performance (Carless and De Paola, 2000), was high in these higher performing teams. The model of shared knowledge of cases also fitted very much with this sense of shared knowledge and responsibility, serving both a pragmatic role (in terms of accounting for staff leave, for example), and a supportive role (in terms of reducing feelings of isolation in dealing with difficult cases or issues).

The highest performing MAU teams exhibited different characteristics from the lowest in a number of areas, as we have seen. To be able to extrapolate from the CYPRESS findings, the same analyses and comparison must be undertaken with a much larger sample of teams, and the relationship with outcome explored.

Observations on the nature of MAU teams. During the interviewing process, an issue emerged which had not been considered during the measure development phase; a key distinction between different MAU teams was that some were services with which involvement was voluntary for all families, and others were

statutory services, that is to say those with which families or young people were compulsorily engaged, as a result of child protection or youth justice legislation, while some worked with both voluntary and compulsory cases. Staff spoke of the differences in terms of engagement in light of whether a young person or family's involvement was voluntary or otherwise. Clearly the degree to which a young person or parent/carer feels invested in the work they are engaged in, and how much this work can be said to be collaborative, are likely to be impacted upon by whether they are engaging by choice or compulsion, and whether the role of the therapist is seen as a helping role, a legalistic role, or a combination of the two (see, for example, Trotter, 1999). It follows then that CYPRESS scores are likely to reflect this, and account must be taken of this feature of services, if CYPRESS is to be applicable across contexts.

The question arose during the research process as to who exactly a team considered to be their client – the young person, the parent/carer(s), or the whole family system. This has implications for how a team scores on items 15 (individualised care) and 16 (family/carer involvement); some workers might consider their practice to be collaborative as regards the young person, to the exclusion of the parent/carer, or indeed vice versa. In MST, the parent/carer is the primary person with whom therapists work; there is no compulsion for the young person to be involved in the meetings. This is quite at odds with the practice in some of the usual services interviewed. The nature of the relationship, and therefore the degree of alliance, between therapist and either young person or parent/carer has wide-ranging implications for the success of therapy, with some research demonstrating improved self and parent reported behaviour as being related to youth (but not parent) alliance with the therapist (Hawley and Weisz, 2005).

Methodological limitations

Development of the measure. Firstly, there are considerations related to the development of the measure. In specifying the areas to be examined, account was taken of existing measures in the field and other relevant documents, but the disparate nature of these meant that there was the potential for an element of the idiosyncratic in the selection of items. It is hoped that this was accounted for through employing three developers to work collaboratively, and using a systematic approach to the stimulus material, but it would be remiss to disregard this possibility, which indeed might exist for any newly developed clinical tool. Further, the planned Delphi process was conducted with only a small number of experts, rather than through a more rigorous and comprehensive approach; this might be said to reduce the confidence with which impartiality or lack of bias can be asserted. It will be valuable to revisit and seek opinion on the domains addressed as part of the process of further development. The raters' dual role as both originators of the measure and raters might have impacted both on the process of scoring and the way in which the usability of the measure was understood. The importance of additional raters conducting inter-rater reliability testing will be valuable not only for establishing this element of the reliability of the measure but also for gaining more objective perspectives on the ease of use.

Sensitivity of the measure. The experience of conducting CYPRESS has allowed the developers to see where there are shortcomings. The measure was shown at times, on some items, to lack sensitivity and scope to reflect accurately the reports of interviewees. This was perhaps most apparent in item 8 (supervision). While teams across the sample scored 5, with only one team scoring 4, the reported experience of supervision showed great variability. The question fails to reflect any

information about the nature or quality of supervision, asking as it does only whether supervision is available to all staff. Staff reported across conditions that opportunities for reflection were highly valued but often came secondary to more action-focussed, didactic supervisory interactions. Clinical supervision has been extensively studied (Ellis, Ladany, Krengel and Schult, 1996) and research demonstrates that supervision is a complex and multifaceted interaction, important for clinician development, in terms of both specific competencies and professional identity, and therapeutic outcomes (see e.g. Falender and Shafranske, 2012; Milne and James, 2000; Milne and Reiser, 2012; Rønnestad and Skovholt, 2003), and while it is uncommon to find examples of health or social care teams which report no supervision, the quality of the supervision, and therefore the practice it supports, might vary dramatically. Items 7 (meetings) and 19 (risk and child protection) similarly demonstrated this shortcoming, giving credit for the presence of a feature but not distinguishing between services in terms of quality. Quality of provision is clearly essential to effective services and this must be accurately reflected in measures of service delivery.

Measure design. Having become very familiar with the use of CYPRESS, practical issues in terms of its use have become apparent, including the need for the layout to be revised so that information is as compact as possible, with the numbers associated with each rating to be visible at the top of every page, if not in every response box. Further, ease of use must be considered. While a manual is available for raters to use alongside the score sheet, the reality might be that to use such a manual whilst simultaneously recording scores and taking notes, without the interview becoming unnecessarily arduous, is too difficult.

Potential for bias in interview. The interview process itself cannot be said to be free of potential hazards and biases. Interviewees in the MST condition often knew the first rater (SB) in the context of the START trial, and it is possible that this prior knowledge, both of the rater in the case of the participants and the reverse, impacted on the reports given by interviewees and the scores given by the rater. Further, participants in both conditions might have felt concerned, despite reassurances of confidentiality and that there would be no individualised reporting, about the implications of voicing negative opinions of the services of which they were part, or of their managers. It is hard to know the degree to which interviewees might have felt compelled to give positive reports of their services, and their relationship with the service (and possibly with the START trial) would likely have impacted on this. Further considerations in terms of how scores might have been influenced relates to the practical question of interview order, both in terms of whether therapists or supervisors were interviewed first, and in terms of whether the MST or MAU team in a particular locality was interviewed first. Those teams which did not make available the CYPRESS pre-interview information form prior to the interview might also have been experienced differently by the raters. Further, teams in which it was possible to interview only one therapist might have been rated differently, both because of the perceptions of the raters of the team, and due to the multi-informant approach not being exercised.

Applicability

We should consider the measure's applicability in the real world – more work is needed to understand if there is an appetite for a measure which can be used across interventions, and by whom it might be used. It might be that organisations in fact have a preference for measures which are specific to a given therapy, but perhaps

incorporate service delivery aspects as well as issues of treatment adherence. CYPRESS would have to be tested extensively to understand whether in practice it is possible or desirable to use a measure which is both pan-model and concerned with service delivery.

Conclusions

The utilisation of CYPRESS in this context and the finding that it can distinguish between teams is valuable. However, this gives us only half the picture; if CYPRESS is to be considered a truly useful tool, the degree to which scores on the measure are related to outcome is of utmost importance. Organisations will only be committed to evaluating and improving their service delivery frameworks if they are demonstrated to impact on the outcomes achieved by teams (cf. Department of Health, 2010).

CYPRESS evaluates fidelity to a set of principles, conceptualised in terms of 20 key areas of functioning, considered to be important for the delivery of complex interventions for children and young people. Certainly without further rigorous examination of the measure and its properties, it cannot be assumed that CYPRESS relates to outcomes or has value in assessing services. At this stage, CYPRESS can only measure what it is explicitly concerned with, and any extrapolation from it should be cautious; we cannot 'rate' teams as such until such a time that we can examine the relationship between these service delivery elements and treatment outcomes. However, the application of CYPRESS in this context indicates that it is possible to implement a fidelity measure which has utility across interventions, and which allows services to be characterised in terms of the service delivery frameworks by which are they supported.

References

British Psychological Society (2010) *Continuing Professional Development Guidelines*. Leicester: British Psychological Society.

Butler, S., Baruch, G., Hickey, N. and Fonagy, P. (2011) A Randomized Controlled Trial of Multisystemic Therapy and a Statutory Therapeutic Intervention for Young Offenders. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*, 12, pp.1220-1235.

Carless, S.A. and De Paola, C. (2000) The measurement of cohesion in work teams. *Small Group Research, 31*, 1, pp.71-88.

Curtis, N.M., Ronan, K.R. and Borduin, C.M. (2004) Multisystemic Treatment: A Meta-Analysis of Outcome Studies. *Journal of Family Psychology, 18*, 3, pp.411-419.

Damschroder, L.J., Aron, D.C., Keith, R.E., Kirsh, S.R., Alexander, J.A, and Lowery, J.C. (2009) Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science, 4*, 50 [online]. Available at: http://www.implementationscience.com/content/4/1/50 . Accessed 13 June 2012.

Department for Education and Skills (2003) *Every Child Matters*. London: The Stationery Office.

Department of Health (2008) *High Quality Care for All*. London: The Stationery Office.

Department of Health (2010) *Equity and excellence: Liberating the NHS*. London: The Stationery Office.

Ellis, M.V., Ladany, N., Krengel, M. and Schult, D. (1996) Clinical Supervision Research From 1981-1993: A Methodological Critique. *Journal of Counseling Psychology, 43*, 1, pp.35-50.

Falender, C.A. and Shafranske, E.P. (2012) The Importance of Competency-based Clinical Supervision and Training in the Twenty-first Century: Why Bother? *Journal of Contemporary Psychotherapy, 42*, 3, pp.129-137.

Ferlie, E.B. and Shortell, S.M. (2001) Improving the Quality of Health Care in the United Kingdom and the United States: A Framework for Change. *The Milbank Quarterly, 79*, 2, pp.218-315.

Garland, A.F., Brookman-Frazee, L., Hurlburt, M.S., Accurso, E.C., Zoffness, R.J., Haine-Schlagel, R. and Ganger, W. (2010) Mental Health Care for Children With Disruptive Behavior Problems: A View Inside Therapists' Offices. *Psychiatric Services, 61*, 8, pp.788-795.

Garland, A.F., Hurlburt, M.S., Brookman-Frazee, L., Taylor, R.M. and Accurso, E.C. (2010) Methodological Challenges of Characterizing Usual Care Psychotherapeutic Practice. *Administration and Policy in Mental Health, 37*, pp.208-220.

Ghauri, P. (2004) Designing and Conducting Case Studies in International Business Research. In R. Marschan-Piekkari and C.A. Welch (Eds.) *Handbook of Qualitative Research Methods for International Business* (pp.109-124). Cheltenham: Edward Elgar Publishing Ltd.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P. and Kyriakidou, O. (2004) Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations. *The Milbank Quarterly, 82*, 4, pp.581-629.

Hawley, K.M. and Weisz, J.R. (2005) Youth Versus Parent Working Alliance in Usual Clinical Care: Distinctive Associations with Retention, Satisfaction and Treatment Outcome. *Journal of Clinical Child and Adolescent Psychology, 34,* 1, pp.117-128.

Health and Social Care Act 2012. London: The Stationery Office [online]. Available at: http://www.legislation.gov.uk/ukpga/2012/7/contents/enacted . Accessed 7 June 2012.

Henggeler, S.W. (1999) Multisystemic therapy: An overview of clinical procedures, outcomes, and policy implications. *Child Psychology and Psychiatry Review, 4*, pp.2–10.

Henggeler, S. W. and Borduin, C. M. (1990) *Family therapy and beyond: A multisystemic approach to treating the behaviour problems of children and adolescents.* Pacific Grove, California: Brooks Cole.

Henggeler, S.W. and Borduin, C.M. (1992) *Multisystemic Therapy Adherence Scales* [Unpublished instrument]. Charleston, South Carolina: Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina.

Hoagwood, K. and Kolko, D.J. (2009) Introduction to the Special Section on Practice Contexts: A Glimpse into the Nether World of Public Mental Health Services for Children and Families [Editorial]. *Administration and Policy in Mental Health, 36*, pp.35-36.

Horwath, J. and Morrison, T. (2007) Collaboration, integration and change in children's services: Critical issues and key ingredients. *Child Abuse and Neglect, 31,* 1, pp.55-69.

Hudson, B., Hardy, B., Henwood, M., and Wistow, G. (1999) In Pursuit of Inter-Agency Collaboration In The Public Sector. *Public Management: An International Journal of Research and Theory, 1*, 2, pp.235-260.

Imison, C., Buchan, J. and Xavier, S. (2009) *NHS Workforce Planning: Limitations and Possibilities*. London: The King's Fund.

In-Albon, T. and Schneider, S. (2007) Psychotherapy of Childhood Anxiety Disorders: A Meta-Analysis. *Psychotherapy and Psychosomatics, 76*, pp.15-24.

James, A.A.C.J., Soler, A. and Weatherall, R.R.W. (2005) Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews 2005*, Issue 4. Art. No.: CD004690. DOI: 10.1002/14651858.CD004690.pub2.

Leadbetter, J. (2008) Learning in and for interagency working: making links between practice development and structured reflection. *Learning in Health and Social Care, 7*, 4, pp.198-208.

Leschied, A.W. and Cunningham, A. (2002) *Seeking effective interventions for young offenders: Interim results of a four-year randomized study of Multisystemic Therapy in Ontario, Canada* [online]. Available at: http://www.lfcc.on.ca/seeking.html . Accessed 20 September 2010.

Milne, D. and James, I. (2000) A systematic review of effective cognitive-behavioural supervision. *British Journal of Clinical Psychology, 39*, 2, pp.111-127.

Milne, D. and Reiser, R.P. (2012) A Rationale for Evidence-Based Clinical Supervision. *Journal of Contemporary Psychotherapy, 42*, 3, pp.139-149.

Mueser, K.T., Noordsy, D.L., Drake, R.E. and Fox, L. (2003) *Integrated Treatment for Dual Disorders: A Guide to Effective Practice*. New York: Guilford Press.

National Institute for Health and Clinical Excellence (2006) *Parent-training/education programmes in the management of children with conduct disorders* [TA102]. London: National Institute for Health and Clinical Excellence.

Rønnestad, M.H. and Skovholt, T.M. (2003) The Journey of the Counselor and Therapist: Research Findings and Perspectives on Professional Development. *Journal of Career Development, 30*, 1, pp.5-44.

Rose, J. (2011) Dilemmas of Inter-Professional Collaboration: Can they be Resolved? *Children and Society, 25*, 2, pp.151-163.

Roth, A.D., Calder, F. and Pilling, S. (2011) *A competence framework for Child and Adolescent Mental Health* Services [online]. Available at: http://www.ucl.ac.uk/clinical-psychology/CORE/child-adolescent-competences/CAMHS%20Competences%20Framework_V1%20(2).pdf . Accessed 18 June 2012.

Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B. and Richardson, W.S. (1996) Evidence based medicine: what it is and what it isn't [Editorial]. *British Medical Journal, 312*, pp.71-72.

Schoenwald, S.K. (1998) *Multisystemic Therapy Consultation Manual*. Charleston, South Carolina: MST Institute.

Schoenwald, S.K. (2008) Toward Evidence-Based Transport of Evidence-Based Treatments: MST as an Example. *Journal of Child and Adolescent Substance Abuse*, *17*, 3, pp.69-91.

Schoenwald, S.K., Henggeler, S.W. and Edwards, D.L. (1998). *MST Supervisor Adherence Measure* [Unpublished instrument]. Charleston, South Carolina: MST Institute.

Sundell, K., Hanson, K., Löfholm, C.A., Olsson, T., Gustle, L.-H. and 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*, 3, pp.550–560.

Teague, G.B., Bond, G.R. and Drake, R.E. (1998) Program fidelity in assertive community treatment: development and use of a measure. *American Journal of Orthopsychiatry, 68*, pp.216-232.

Trotter, C. (1999) *Working with Involuntary Clients: A guide to practice*. London: SAGE Publications Ltd.

Utting, D., Monteiro, H. and Ghate, D. (2007) *Interventions for Children at Risk of Developing Antisocial Personality Disorder: Report to the Department of Health and Prime Minister's Strategy Unit*. London: Policy Research Bureau with Department of Health and the Cabinet Office.

Part 3: Critical Appraisal

This critical appraisal considers further some of the issues raised in the empirical paper, and the implications for future research, and reflects on some of the challenges and questions raised.

Methodological issues

The possibility for different biases to have influenced the development of the Children and Young People – Resources, Evaluation and Systems Schedule (CYPRESS) should not be underestimated. Firstly, as we have noted, CYPRESS was developed in the context of a trial looking at the effectiveness of multisystemic therapy (MST) in the UK. While efforts were made to use a systematic and broad approach to understanding the appropriate structure and content for the measure, the potential for the developers of CYPRESS to be shaped by the well-developed practice and ethos of MST, in terms of what features of service delivery were valued, existed. Consequently, wider examination by professionals not involved in the START trial or in MST in other capacities will be beneficial in establishing the degree to which CYPRESS can be said to be balanced and representative of the service delivery frameworks underpinning complex interventions for children and young people in general, rather than being biased towards MST, in terms of the features that it rewards. This is not to say that MST might not continue to score highly on CYPRESS, as we know that MST is a highly developed, evidence-based intervention for young people and their families, which emphasises service delivery elements which assure quality, such as those measured in CYPRESS. Indeed, these qualities have been adopted by many other services because they are seen to be valuable. Whatever its origins or influences, the measure must be balanced and offer the possibility for non-MST teams to score highly. We have already seen that this seems to be possible based on the 16 teams interviewed; the joint (with one other)

second highest scoring site was a management as usual (MAU) service – this would need to be replicated on a larger scale, for the purposes of establishing the validity of this finding and others.

While the development of CYPRESS was informed by existing measures and literature on the delivery of complex interventions (see, for example, Utting, Monteiro and Ghate, 2007), further exploration is needed to help us to understand whether CYPRESS measures the appropriate constructs. It is hoped that CYPRESS captures key tenets of practice at the service delivery framework level. In terms of establishing whether the relevant constructs are being measured, the literature review brought together existing measures and relevant articles which could inform us about the structure and content appropriate to the measure. While every effort was made to ensure this was done in a rigorous manner, it was nevertheless a creative process at times; there are many ways in which one might set about designing such a measure and this was one such way, there may very well be superior alternatives not explored here.

Once we have access to the outcomes and are able to compare them with the CYPRESS scores, we will be able to understand if the constructs measured in CYPRESS relate to outcomes and in what way. Our assumption is that those services which demonstrated the most evolved service delivery frameworks will be the same services which achieve the best outcomes with young people; when we are able to examine this relationship we will be able to more definitively establish whether CYPRESS, or indeed other measures of this nature, are valuable and viable.

There are also significant issues related to the administration and scoring of CYPRESS. Firstly, it will be an important phase in the development and evaluation of CYPRESS to establish the inter-rater reliability and other psychometric properties,

which unfortunately were not assessed within the time frame necessary to be included here. Secondly, it will be necessary to examine the degree to which CYPRESS is useable by those other than the developers, not only in terms of whether non-developers using the measure tend to reach similar scores to one another during the independent scoring phase, and whether they subsequently reach similar scores to the developers, as noted above, but also in terms of ease of use and the acceptability to other raters.

In considering the administration of the CYPRESS interviews, the fact that a number of respondents in the MST teams knew the first rater (SB) must also be borne in mind. In this respect, it is possible that these prior relationships between one of the raters and the MST teams influenced the manner in which therapists and supervisors responded to the questions asked of them, and also the rating of the teams by the rater. However, the second rater (CG) was not known to the teams prior to the initiation of the CYPRESS procedure, and yet during independent rating tended to score the teams very similarly to the first. This might be taken to mean that prior knowledge was not an important factor, but we should nonetheless be mindful of this potential confounding variable in understanding the results.

The nature of the MAU teams, and the manner in which they were identified, must be borne in mind. The MAU services interviewed here represent only some of the usual services utilised by children and young people in the trial areas, and as such cannot be taken to represent MAU in its entirety. Further, in this context MAU teams were treated as if equivalent to MST teams, however, indeed our analysis has shown this, they might be very different services by their nature. It is possible that some of the services seem disparate or less well functioning than their MST counterparts because of the way in which they have been defined. While MST teams are uniquely

well defined in terms of their service parameters and boundaries (see Henggeler, 1999), other services might share team managers and other features of the service, yet represent distinct elements of a wider service, and this might not be reflected accurately by the CYPRESS process. What we see from the experience of carrying out the research and from analysis of the findings, is that MAU can be different and varied, and as such the extent to which they can be considered exact counterparts of the MST teams cannot be taken for granted. It seems likely that the true equivalence or otherwise of the teams under examination has the potential to influence the scores on CYPRESS, and this warrants further consideration. It might be that if CYPRESS is to be used as a comparison tool, a test of whether two services can be considered equivalent would also have to be established.

The identification of MAU, and securing their agreement to take part in the CYPRESS procedure in particular, was often challenging. MST teams varied greatly in the sense in which they felt able or willing to identify the appropriate MAU team. This was not an exact science, and while centrally collected data held by the trial coordinator was available to use in this process, a number of instances of randomisation to MAU were not accompanied by details of the specific MAU team. Further, we cannot have certainty that the MAU teams identified were not inclined towards services with which the relevant MST team had a better relationship, and this is another potential source of bias.

A key consideration has been how best to analyse and understand the data collected. Rich information and varied opinions came out of the interviews, and much of this went beyond what we see reflected in the CYPRESS scores. On reflection I think a valuable addition to the work would be a qualitative element to the characterisation of MAU, using a method such as interpretative

phenomenological analysis (IPA; see Smith and Osborn, 2008). This consideration emerged after the interviews were underway and it is a shortcoming of the research that this was not built into the original design. A qualitative approach such as this might not be appropriate for routine use with CYPRESS in future, but it might have helped to build a richer picture of the MAU services we were attempting to characterise, in the context of the START trial and the development of this new measure. A formal qualitative analysis would offer a more precise lens through which to examine the similarities and differences between the teams interviewed. Supervision is one example of an area in which the verbal reports of the therapists, in particular, but also of the supervisors, indicated experiences of supervision which were not reflected in the CYPRESS scores. There was a recurring theme among the MST therapists, for example, around the highly structured and action-oriented nature of supervision in MST, which seemed to preclude more reflective processes. It might also be interesting to explore themes in the nature of the language used by staff to speak about their role and the work of the team, as, anecdotally, there was much variation in the tone and content of their verbal reports.

Implications of the study and future research

There is an ever growing focus in modern health and social care on evidence-based practice and maximising the benefits of interventions so as to achieve the best outcomes possible and by the most cost-effective means. There is a contemporaneous development of understanding as regards how we can best measure these elements.

The examples of this trend are various. The Children and Young People's Improving Access to Psychological Therapies (CYP IAPT) programme emphasises strongly the value of session by session outcome monitoring as a fundamental feature of practice (see for example the CYP IAPT guidance on the use of outcome measures

in routine clinical practice; Law, 2012). Further, and in line with what has been argued, this guidance comments on the importance of structures which support such monitoring and allow it to be embedded in practice.

Considered a gold standard of evidence-based practice (EBP), and having shown positive results in early outcome research in the UK (Butler, Baruch, Hickey and Fonagy, 2011), MST teams are now being rolled out across the country as a result of government investment (Department for Education, 2012). Meanwhile, Sure Start centres are subject to a reform programme which sees an emphasis on EBP and payment by results (Department for Education, 2010). The effects of recent budget cuts on Sure Start centres have perhaps had more of an impact in practice than has the intended reform programme, but the stated intention nonetheless demonstrated the trend for making accountability, and the use of evidence to drive practice, central issues.

These are just some examples of the way in which EBP, and routine measurement of both what practitioners and organisations offer children and young people, and the associated outcomes, are increasingly taking centre stage at the policy and clinical level.

CYPRESS is one example of efforts to bring together thinking about the way in which services function in a manner which allows organisations to make the best use of the expertise held within them. It will be necessary to establish whether services working with children and young people identify with the need to be able to measure services in this way; again an understanding, once it is possible, of the relationship between service delivery frameworks and outcomes will be valuable here. Given the importance of EBP and the development of frameworks to support implementation, it will be valuable to establish how CYPRESS might be used; would

it be used by researchers or those conducting audit, that is to say people external to the team, or could it be used as an internal self-monitoring tool for organisations wishing to better characterise and measure their existing service delivery frameworks? Indeed, it seems possible that it could be used for a variety of purposes.

As has been noted, the first step in terms of future research would be to administer CYPRESS very much more widely and gather a sample of CYPRESS data from a much larger group of teams. This might initially be restricted to teams working with children and young people with conduct problems and associated difficulties, as seen in START. However, it would be interesting to explore whether CYPRESS could in fact be applied in other settings serving children and young people, and their families, and indeed beyond this to other areas of health and social care altogether. The principle that CYPRESS aims to capture service delivery elements which support interventions, rather than features of a specific therapy, means that in theory this should be possible. This could offer more rigorous examination of the scores across services, providing as it would a sample size which would allow more robust statistical analysis. As noted, the measure must be examined in terms of its validity and inter-rater reliability, before it can be considered appropriate for wider use.

It will be important to consider what increased knowledge of both MST and MAU has on understanding what is unique about MST, a time and money intensive resource, as compared with usual services. This underlines the value of a measure such as CYPRESS – if we can characterise usual services more effectively, we can understand the respective value and roles of both MST and MAU.

Reflections on the process

This area of investigation presented conceptual challenges which had to be addressed in order to progress with the work, yet which I continued to consider throughout the research and during the write up process. The terminology around fidelity is in and of itself complex and used in different ways, at times subtly so. I found that reading around the subject was illuminating but at times also served to raise more questions than it answered. While this serves the function of creating a richer background from which to work, it can be challenging nonetheless.

This endeavour has been hugely interesting and has challenged me to think in novel ways about how we understand services and the means by which we evaluate them. Conversations with the staff who were the respondents in the CYPRESS process indicated that for many the consideration of these service delivery frameworks was also novel for them, yet extremely interesting.

There were practical and bureaucratic obstacles to progress at times. It was in some instances difficult to establish the participation of MAU teams, and I experienced a range of responses. Some of the MAU services seemed not wholly keen on being interviewed, seeming to feel they were being judged (and it would be disingenuous to suggest that this was not in part the case); some teams seemed to feel that taking part was a burden in the context of busy schedules – this was my sense of the reactions of some teams, but of course this is only an anecdotal impression and was not established in any systematic way. The fact that three MAU services gave us no data beforehand might be considered a symptom of these difficulties. Securing the agreement of one MAU team to represent each area was more challenging than anticipated. Indeed, in the case of one location, it was not possible to establish a team which might be willing to take part for some months. When a team was eventually

identified, it took several weeks to establish direct contact (previously this had been through the relevant MST supervisor), and this subsequently stopped, with my efforts at contact going unanswered, meaning that this location had to be excluded from the CYPRESS process. While frustrating, this was an interesting insight into some of the experiences of the MST team in that area, where relationships with existing services, and the role of MST, were very much less established than in other parts of the country.

It should perhaps not be too surprising that it was more challenging to engage the MAU services in the process, as they were involved in the START trial at arm's length; while the MST teams were in place as a result of the research and associated funding, the other teams were by definition pre-existing services which happened to be in areas where MST was set up as part of the trial. It was interesting to note the differing opinions of MST; in some MAU services staff were positive about MST and its role in the area, considering it a useful adjunct to what was already available. In other areas, there was a sense that MST was considered to some extent a (well-resourced) rival, and thought of less highly.

One broader reflection relates to the question of the degree to which a standardised measure of service delivery frameworks is possible and the degree to which it loses sight of the people involved – those delivering therapeutic interventions and the clients served by them. While it might be possible to create a measure which appears valid, it would be important to go back to the service and ask if the characterisation derived from the use of the measure is an accurate reflection of what the service really looks like. Do practitioners and organisations work in the arguably rigid ways that a measure such as CYPRESS might suggest, and if so, should they? It might be argued that the focus on evidence-based practice, increased

use of manuals, and outcome measurement fails to recognise the value of flexibility in therapy, at the individual and the organisational level, of treating each client, indeed each team or organisation, as individual, of a human approach to human problems.

Apparent in conversation with the majority of the MST staff interviewed was a steadfast belief in and dedication to MST and their role. The nature of the post is such that one imagines it would not be possible to do the job without this strength of belief, requiring as it does a flexible approach to working hours and practices, and a willingness to routinely work in challenging situations where they are exposed to high levels of emotion and might be considered to be at heightened risk of harm. The impact of this degree of belief in MST should perhaps be considered in understanding our findings. Firstly, we should consider the possibility that staff – in the MST team in particular, given this observation as regards the level of commitment, but also in the MAU teams – might be inclined to report on their team more favourably if they feel strongly allied with the service or the intervention. In the case of all the teams interviewed the information gathered is mediated by the staff members' opinions of and feelings about their clinical work, role, team and organisation, which in turn might be influenced by a number of factors. Secondly, the degree of commitment shown by an employee to the organisation or service in which they work might have more complex implications for understanding our results. If a member of staff is highly committed to their organisation they are more likely to engage in behaviours which are in line with relevant organisational or team goals (see Meyer and Herscovitch, 2001); while in-depth consideration of the role of employee commitment is beyond our scope here, it seems possible that performance on CYPRESS could be mediated by this. Indeed, the individual characteristics of the

particular members of staff interviewed also have the potential to influence their responses and the raters' scores of the team. The degree to which the respondents might or might not be representative of the wider team should also be considered.

The process of reporting on the CYPRESS project has led me to reflect on the wider social and political context in which all that has been discussed is located.

Different countries have differing approaches to children and young people who commit crime or engage in antisocial behaviour. This is exemplified by different ages of criminal responsibility, and differing legislative responses. Consideration of this is important not only to broadly contextualise the experiences of the young people in the START trial, but also to understand the background to the usual services with which MST is compared. The UK might be imagined to adopt a more welfare-based, as opposed to punitive, approach to such young people as compared with the US, where MST has been demonstrated to be highly effective; this was certainly my assumption at the outset. While the proportion of under-18s in custody in the US is indeed significantly higher than in the UK and the rest of Europe (see Hazel, 2008), consideration of the literature on responses to criminality, both generally and in respect of young people, leads us to examine this general assumption more closely. Garland (2001) argues convincingly that the latter part of the last century saw a spread of what he describes as US ideals to the UK context, with the 'penal welfarism' that proliferated in the UK and elsewhere from the late nineteenth century until the 1970s being reversed. He argues that, perhaps in response to the increased personal and social freedoms of the 1960s and 70s, the years following these decades saw increased levels of civic mistrust. This in turn is said to have led to a return to more intensive regulatory regimes and an increased focus on policing, penalty and prevention of crime, in both the US and the UK.

If we accept Garland's thesis, we should consider how this relates to the treatment of children and young people. While a given approach to responding to crime or antisocial behaviour might appear to be progressive and child-centred, this overlooks the ever present threat (of custody) that hangs over any efforts to engage with young people who are involved in criminal or otherwise antisocial behaviour, against the increasingly punitive and regulated social backdrop Garland describes. Indeed, the intensity with which MST therapists engage with families, and the preference for disengaging young people from antisocial peers, might be argued to represent a panopticistic state of affairs, in which surveillance (or the possibility of it) and isolation from one's usual network of peers are used as a means of control. There is an argument that interventions of this nature at one level constitute complicity with social control, meaning that their growth might be argued to represent providers of services being co-opted into a regulatory function.

Muncie (2008) addresses the question of the 'punitive turn' with specific reference to juvenile justice, linking this to the rise of a neoconservative agenda as regards crime and punishment in the UK and Europe more generally. Muncie's view is a little more optimistic than Garland's; he comments that while a general cultural shift towards more punitive ways of responding is evident, examination of different countries demonstrates that this is mediated by national and local practitioner values, which might offer progressive alternatives to the most punitive responses to young people who engage in socially unacceptable behaviour, both criminal and otherwise.

A more extensive consideration of these issues is beyond the scope of this appraisal, but it should be noted that the ways in which antisocial behaviour and collective responses to it are conceptualised has implications for our understanding of both MST and MAU.

In conclusion, the development of a measure which enables further exploration of not only MAU but also MST, in terms of the service delivery frameworks which underpin them, is a valuable addition to this area of research. CYPRESS appears to allow us to distinguish between different services in terms of those service delivery frameworks, but needs further work to understand its validity, reliability, and real world applicability on a large scale. All research in this area should be conducted with a sensitive consideration of the wider issues at play in the arena of crime, youth antisocial behaviour and the associated societal responses, and the social problems and cultural phenomena which underlie these matters, as well as the underlying discourses by which all of these issues are informed.

References

Butler, S., Baruch, G., Hickey, N. and Fonagy, P. (2011) A Randomized Controlled Trial of Multisystemic Therapy and a Statutory Therapeutic Intervention for Young Offenders. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*, 12, pp.1220-1235.

Department for Education (2010) *Business Plan 2011-2015* [online]. Available at: http://media.education.gov.uk/assets/files/pdf/d/department%20for%20education%20business%20plan.pdf . Accessed 5 June 2012.

Department for Education (2012) *Business Plan 2012-2015* [online]. Available at http://media.education.gov.uk/assets/files/pdf/d/department%20for%20education%20business%20plan%202012%202015.pdf . Accessed 5 June 2012.

Garland, D. (2001) *The Culture of Control: Crime and Social Order in Contemporary Society*. Oxford: Oxford University Press.

Hazel, N. (2008) *Cross-national comparison of youth justice*. London: Youth Justice Board.

Law, D. (Ed.) (2012) *A Practical Guide to Using Service User Feedback and Outcome Tools to Inform Clinical Practice in Child and Adolescent Mental Health* [online]. Available at: http://www.iapt.nhs.uk/silo/files/a-practical-guide-to-using-service-user-feedback--outcome-tools-.pdf . Accessed 5 June 2012.

Meyer, J.P. and Herscovitch, L. (2001) Commitment in the workplace: Toward a general model. *Human Resource Management Review, 11*, 3, pp.299-326.

Muncie, J. (2008) The 'punitive' turn in juvenile justice: Cultures of control and rights compliance in western Europe and the USA. *Youth Justice, 8*, 2, pp.107-121.

Smith, J.A. and Osborn, M. (2008) Interpretative Phenomenological Analysis. In J.A. Smith (Ed.) *Qualitative Psychology: A Practical Guide to Research Methods* (2nd ed.) (pp.53-80). London: SAGE Publications Ltd.

Utting, D., Monteiro, H. and Ghate, D. (2007) *Interventions for Children at Risk of Developing Antisocial Personality Disorder: Report to the Department of Health and Prime Minister's Strategy Unit*. London: Policy Research Bureau with Department of Health and the Cabinet Office.

Appendix A: MST Therapist Adherence Measure

Multisystemic Therapy Institute
Therapist Adherence Measure- Revised (TAM-R)

Current Supervisor: _____
Therapist Code: _____
Family Code: _____
Respondent _____ Caregiver _____
Date Form Completed (mm-dd-yyyy): _____

Please answer the following two questions:

How many times has the therapist met with your family within the last week? _____
If zero, then when did the therapist last see anyone in the family (mm/dd/yyyy)? _____

Regarding your last 2-3 sessions:

	Did Not Respond	Not at All	A Little	Some	Pretty Much	Very Much
1. The therapist tried to understand how my family's problems all fit together.						
2. My family and the therapist worked together effectively.						
3. My family knew exactly which problems we were working on.						
4.The therapist recommended that family members do specific things to solve our problems.						
5. The therapist's recommendations required family members to work on our problems almost every day.						
6. The therapist understood what is good about our family.						
7. My family and the therapist had similar ideas about ways to solve problems.						
8. The therapist tried to change some ways that family members interact with each other.						
9. The therapist tried to change some ways that family members interact with people outside the family.						
10. My family and the therapist were honest and straightforward with each other.						
11. The therapist's recommendations should help the children to mature.						

	Did Not Respond	Not at All	A Little	Some	Pretty Much	Very Much
12. Family members and the therapist agreed upon the goals of the session.						
13. My family talked with the therapist about how well we followed her/his recommendations from the previous session.						
14. My family talked with the therapist about the success (or lack of success) of her/his recommendations from the previous session.						
15. We got much accomplished during the therapy session.						
16. My family was sure about the direction of treatment.						
17. The therapist's recommendations made good use of our family's strengths.						
18. My family accepted that part of the therapist's job is to help us change certain things about our family.						
19. The therapist's recommendations should help family members to become more responsible.						
20. The therapist talked to family members in a way we could understand.						
21. Our family agreed with the therapist about the goals of treatment.						
22. The therapist checked to see whether homework was completed from the last session.						
23. The therapist did whatever it took to help our family with tough situations.						
24. The therapist helped us to enforce rules for the child.						
25. The therapist helped family members talk with each other to solve problems.						

	Did Not Respond	Not an Issue	Not at All	A Little	Some	Pretty Much	Very Much
26. The therapist helped us keep our child from hanging around with troublesome friends.							
27. The therapist helped us improve our child's behavior at school.							
28. The therapist helped us get our child to stay in school every day.							

Appendix B: MST Supervisor Adherence Measure

SAM Form

(Transportability version)

Form ID
Organization
Team
Supervisor Name
Therapist
Current Consultant/System Supervisor
Date Form Completed (mm/dd/yyyy)

Please consider your supervision sessions over the past two months as you complete the following items.

1. When the supervisor recommended changes in my course of action, the rationale for the recommendation was described in terms of one or more of the MST principles.	Never	Rarely	Sometimes	Often	Almost Always
2. You could tell that the supervisor was in charge of the sessions.	Never	Rarely	Sometimes	Often	Almost Always
3. Team members took a long time to describe the details of cases before the supervisor spoke.	Never	Rarely	Sometimes	Often	Almost Always
4. The supervisor asked clinicians for evidence to support their hypotheses about the causes of problems targeted for change or of barriers to intervention success.	Never	Rarely	Sometimes	Often	Almost Always
5. The supervisor asked clinicians how descriptions of this week's case developments pertained to identification of barriers to success.	Never	Rarely	Sometimes	Often	Almost Always
6. When clinicians talked about events in the distant past, the supervisor recommended that current interactions within the family and between family members and others be examined first.	Never	Rarely	Sometimes	Often	Almost Always
7. When clinicians reported on a variety of interventions tried during the week, the supervisor asked for clarification regarding which intermediary goals the interventions aimed to address.	Never	Rarely	Sometimes	Often	Almost Always
8. The supervisor followed up on recommendations made in previous supervision sessions.	Never	Rarely	Sometimes	Often	Almost Always
9. When interventions were not successful, discussion focused on identifying the barriers to success and actions the clinician should take to overcome them.	Never	Rarely	Sometimes	Often	Almost Always
10. I have the skills to implement all of the recommendations made in supervision.	Never	Rarely	Sometimes	Often	Almost Always
11. Interventions that were discussed targeted sequences of interaction between family members.	Never	Rarely	Sometimes	Often	Almost Always
12. Clinicians received positive feedback during the sessions.	Never	Rarely	Sometimes	Often	Almost Always
13. The supervisor asked clinicians how descriptions and questions about case developments pertained to "fit" assessment.	Never	Rarely	Sometimes	Often	Almost Always
14. It was easy for team members to acknowledge frustrations, mistakes, and failures.	Never	Rarely	Sometimes	Often	Almost Always
15. When a clinician presented information about events that transpired during the week, the supervisor asked the clinician and team to clarify the relevance of the information to one or more steps of the analytical process.	Never	Rarely	Sometimes	Often	Almost Always
16. Weekly case summaries were referred to during the discussion of cases.	Never	Rarely	Sometimes	Often	Almost Always

17. Interventions that were discussed targeted sequences of interaction between family members and individuals at school, in the child's peer group, or in the neighborhood.	Never	Rarely	Sometimes	Often	Almost Always
18. When an intervention was only partially successful, the supervisor asked questions to determine whether the clinician had adequately and completely implemented the intervention.	Never	Rarely	Sometimes	Often	Almost Always
19. We spent more time discussing cases in which progress was limited.	Never	Rarely	Sometimes	Often	Almost Always
20. The supervisor referred to specific MST principles while discussing cases.	Never	Rarely	Sometimes	Often	Almost Always
21. The supervisor made a note of case-specific recommendations.	Never	Rarely	Sometimes	Often	Almost Always
22. When new areas were targeted for intervention, the supervisor encouraged the clinician to articulate new intermediary goals accordingly.	Never	Rarely	Sometimes	Often	Almost Always
23. Outcomes were described in observable and measurable terms.	Never	Rarely	Sometimes	Often	Almost Always
24. When clinicians reported plans to meet with teachers, neighbors, or officials from other agencies, the supervisor asked what it would take for a caregiver to hold the meeting.	Never	Rarely	Sometimes	Often	Almost Always
25. When clinicians reported that things were going well in a case, the supervisor focused discussion on factors in the natural ecology that were sustaining progress.	Never	Rarely	Sometimes	Often	Almost Always
26. When clinicians reported doing things for family members, the supervisor asked what it would take for family members to do these things for themselves.	Never	Rarely	Sometimes	Often	Almost Always
27. When clinicians reported that they discussed a particular problem with a family, the supervisor asked what plans were put in place to address the problem this week.	Never	Rarely	Sometimes	Often	Almost Always
28. When clinicians described their ideas about the causes of problems, "fit circles" were developed and discussed in session.	Never	Rarely	Sometimes	Often	Almost Always
29. When clinicians talked about events in the distant past, the supervisor asked for evidence that these events are contributing to a current problem.	Never	Rarely	Sometimes	Often	Almost Always
30. In the past two months, the supervisor and I have discussed the extent to which my case summaries and in-session presentations are consistent with the MST principles and analytic process.	Never	Once	Twice	3-5 Times	Weekly
31. In the past two months, the supervisor and I have set goals for development of my specific competencies in MST.	Never	Once	Twice	3-5 Times	Weekly
32. In the past two months, my supervisor has accompanied me to therapy sessions (i.e., field supervision) OR reviewed audiotapes of my therapy sessions.	This has never happened	This has happened in past, but not in last two months	Once in the last two months	Twice in the last two months (e.g., once a month)	At least twice, and more when necessary
33. In the past two months, I left supervision knowing how to carry out recommended actions.	Never	Once	Twice	3-5 Times	Weekly
34. How skilled do you think your supervisor is in implementing MST interventions?	Not very	Somewhat	Moderately	Very	Extremely

35. How skilled do you think your supervisor is in the treatment modalities used in MST such as cognitive-behavioral therapy?	Not very	Somewhat	Moderately	Very	Extremely
36. How often have you and your supervisor met to develop and monitor a plan to help you increase your knowledge and skill in MST?	Never	Plan developed but never monitored and followed	Plan developed but monitored and followed infrequently	Plan developed but monitored and followed periodically, e.g. once every 3 months	Plan developed but monitored and followed regularly, e.g., once per month

Notes (250 character limit)

Thank you for taking the time to complete this survey.

Appendix C: MST Consultant Adherence Measure

Multisystemic Therapy Institute
Consultant Adherence Measure

Organization	
Team	
Consultant/System Supervisor	
Therapist or Supervisor Name	
Respondent	
Last MST Consultation Date	

For questions 1 - 19, please think about your last MST consultation session

	Never	Rarely	Sometimes	Usually	Almost Always	Always
1.The consultant explained how to implement specific intervention strategies for a case						
2.You could tell the consultant had case-specific ideas about barriers to success and how to overcome them						
3.The consultant helped when the team was “stuck” on some aspect of a case						
4.The consultant described interventions in sufficient detail that clinicians could carry them out						
5.The consultant addressed clinician behaviors that facilitate engagement or treatment progress in specific cases						
6.The consultant was competent at his/her job						
7.The consultant really listened when clinicians talked						
8.The consultant gave positive feedback to clinicians						
9.The consultant conveyed a sense that she/he and the team are “in it together.”						
10.You could tell the consultant had the best interests of the client and clinicians at heart						
11.The consultant gave supportive feedback to clinicians when needed						
12.The consultant referred to specific MST principles when discussing cases						
13.The consultant helped generate a more comprehensive understanding of the “fit” of a problem						
14.The consultant explained what he/she was doing and why						
15.The consultant helped clinicians prioritize problems and intervention targets						
16.The consultant tried to gauge clinician “buy in” to his/her recommendations						
17.Consultation was well structured						
18.The consultant conveyed a “can do” attitude						
19.You could tell the consultant was well prepared						

Multisystemic Therapy Institute
Consultant Adherence Measure

Please answer the following questions about your consultant overall	Not at all	A Little	Somewhat	Quite	Very	Extremely
20. How knowledgeable do you think your consultant is in the theory of MST?						
21. How skilled do you think your consultant is in treatment modalities used in MST such as strategic, structural, behavioral, cognitive-behavioral, therapies?						
22. How skilled do you think your consultant is in implementing MST interventions?						
23. How skilled do you think your consultant is in teaching clinicians to do MST?						

Appendix D: Kumpfer and Alvarado's 13 principles of effective family-focussed interventions

Principles of Effective Family-Focused Interventions

1. Comprehensive multicomponent interventions are more effective in modifying a broader range of risk or protective factors and processes in children than single component programs.

2. Family-focused programs are generally more effective for families with relationship problems than either child-focused or parent-focused programs, particularly if they emphasize family strengths, resilience, and protective processes rather than deficits.

3. Components of effective parent and family programs include addressing strategies for improving family relations, communication, and parental monitoring.

4. Family programs are most enduring in effectiveness if they produce cognitive, affective, and behavioral changes in the ongoing family dynamics and environment.

5. Increased dosage or intensity (25–50 hours) of the intervention is needed with higher risk families with more risk factors and fewer protective factors and processes than low-risk universal families who need only about 5 to 24 hours of intervention.

6. Family programs should be age and developmentally appropriate with new versions taken by parents as their children mature.

7. Addressing developmentally appropriate risk and protective factors or processes at specific times of family need when participants are receptive to change is important.

8. If parents are very dysfunctional, interventions beginning early in the life cycle (i.e., prenatally or early childhood) are more effective.

9. Tailoring the intervention to the cultural traditions of the families improves recruitment, retention, and sometimes outcome effectiveness.

10. High rates of family recruitment and retention (in the range of 80%–85%) are possible with the use of incentives, including food, child care, transportation, rewards for homework completion or attendance, and graduation.

11. The effectiveness of the program is highly tied to the trainer's personal efficacy and confidence, affective characteristics of genuineness, warmth, humor, and empathy, and ability to structure sessions and be directive.

12. Interactive skills training methods (e.g., role plays, active modeling, family practice sessions, homework practice, and videos/CDs of effective and ineffective parenting skills, etc.) versus didactic lecturing increase program effectiveness and client satisfaction particularly with low socioeconomic level parents.

13. Developing a collaborative process whereby clients are empowered to identify their own solutions is also important in developing a supportive relationship and reducing parent resistance and dropout.

Source: Kumpfer and Alvarado, 2003

Appendix E: Zazzali et al.'s adoption, implementation & continuance of evidence based practices semi-structured interview schedules

Semi-Structured Interview Protocol for Clinical Staff
Pilot Study of Barriers & Facilitators to
Functional Family Therapy (FFT) Implementation in NY State

Introduction

This is a voluntary interview and you may choose not to answer particular questions or to end the interview at any time. We would like to ask you some questions about your experience as a clinician in adopting FFT. We are especially interested in hearing about the factors that made this an easy or difficult experience for you. We know from talking with OMH that there are opportunities for improving the implementation of this program. OMH would like to know about factors that will help them create a more effective implementation of this intervention and interventions in the future.

Background information

1. Tell us about your general professional background and role in this organization?
2. How did you happen to get involved in FFT (i.e., self-selection, invitation, delegation)?
3. Are you an original FFT team member or a replacement? Have you participated in externship?
4. How many FFT cases are you expected to carry at any given time? How many do you have right now? If this is not your entire caseload, what does the balance look like (type of case and how many)?

Process of implementation *(stage of adoption, variation in implementation)*

For those organizations that have started or completed FFT training:

5. When were you first trained in FFT and what components of training did you receive?
6. How does this training compare with other types of training (if any) you've received through your organization?

For those organizations that are currently delivering FFT services:

7. What things did you or others in your organization modify as FFT has been implemented (e.g., redesigning programs, or making changes in your facility/center or how people do their work)?
8. Are there any changes that you would like to make in your facility/center that would make implementing FFT easier (e.g., team selection or internal communication about the program)?
9. How hard has FFT been to implement – clinically? - operationally? What types of things have made it more or less difficult?

Organizational context *(norms and attitudes, process of care, resources)*

10. Have you and/or your organization had prior experience implementing other types of evidence based treatment(s) aside from FFT? If so, what were they and did they affect the implementation of FFT at your organization?
11. Do you and your colleagues all feel the same way about the FFT model and its value for your clients?
12. How does FFT compare with other forms of treatment you have done or are currently doing with non-FFT cases? (Probe for whether it is compatible and/or represents a change in practice.)
13. Is there enough staff at your organization with the time and relevant skills to implement and manage FFT (both within and external to the team)?
14. Is there information technology and computer support to implement and manage FFT?
15. What organizational supports are available to you, to enhance your adherence to the model? (Probe for supports to do Engagement & Motivation, to be relentless in overcoming barriers, etc.)

Mechanisms of diffusion *(influence of peers/leaders, change agents, incentives)*

16. What did you know about FFT before its implementation in your organization and what was the source of this information? How has your impression changed since being trained in FFT? (Differentiate between the model, the training process and other operational aspects.)
17. What have you heard about the experiences with FFT programs at other facilities/organizations?
18. Who would you consider “opinion leaders” important to the staff here, and what would you say have been their views of FFT?
19. How would characterize the view of your organization’s management toward FFT? How would you characterize your supervisor’s view?
20. How would you characterize the views of your organization’s clients’ toward FFT? Did the clients play any role in the adoption provcess?
21. Have their been any “champions” of FFT in your organization (i.e., someone who created a vision for the way things could be once FFT is implemented, someone who anticipates barriers to FFT implementation and deals with them, someone who will run interference for the team, etc.)
22. Are there any incentives, financial or otherwise, in place at your organization for performance in adolescent care?
23. Are there other benefits to offering FFT services, such as improved job satisfaction or knowing that your work is making a difference?

Overall assessment and future prospects of program

24. How successful do you feel the FFT program has been so far, or what kind of success do you anticipate?

25. What kinds of client outcomes have you been getting (or anticipate) by using FFT?
26. How have clients reacted to FFT?
27. Has FFT become any easier to do over time? (Probe for clinical care vs. documentation.)
28. What do you see as the prospects of FFT and the ability of sustaining it at your organization?
29. What kinds of things would be helpful for your organization or central OMH and FFT trainers to do in sustaining the program?

Feedback on interview process and questions

30. Do you have any feedback that you would like to share with us about these questions? Are there questions we should be asking that would help us better understand the process by which organizations like yours implement evidence-based practices like FFT?

Semi-Structured Interview Protocol for Agency Lead
Pilot Study of Barriers & Facilitators to
Family Functional Therapy (FFT) Implementation in NY State

Introduction

This is a voluntary interview and you may choose not to answer particular questions or to end the interview at any time. We would like to ask you some questions about your experience as an agency lead in adopting new clinical or service programs, practices, or models. We would like to get your perspective on the adoption process first for programs other than FFT. Then, secondly, we will ask you some questions about your experience with FFT. For our first set of questions, we are especially interested in your experience with new programs, services or models that required training of your staff in the practice or changes in supervision for your staff.

Adoption of New Programs and Program Change Experiences

1a. In the past 5 years, has your agency adopted a new program, service or model?

Yes No

1b. If No new programs, services or models were adopted, why not?

1c. If Yes, what is/are the name(s) of the new program(s), service(s) or model(s)?

1._____

2._____

3._____

4._____

1d. If Yes, which represents the most significant program change, and why? By significant we mean changes in staffing, funding, organizational structures, organizational or program goals, or normal activities.

Thinking about adopting new programs, services, or models in your organization, please answer the following questions.

2. What drives adoption of new programs, services, or models in your organization? (Probe for internal and external factors.)

3. What steps does your organization take in adopting new programs, services or, models (e.g., informing internal and external stakeholders regarding the rationale, implications, priority, what implementation would look like; developing an implementation and support infrastructure, including staff selection and support processes, roles and responsibilities, resource allocation, communications and tracking systems, policies and procedures, outcome evaluation)?

4. What, if any, are the major obstacles that impeded the successful adoption of new programs, services, or models?

5. What are the major factors that make or could make the adoption of new programs, services, or models easier and/or successful?

6. What role do you as agency lead play in the adoption process?

7. What role do the supervisory staff play in the adoption process?

8. What role do clinical staff play in the adoption process?

9. What role do consumers play in the adoption process?

10. What role does OMH play in the adoption process?

11. Out of all of these groups (or others not listed), are there any clear champions (individuals or groups) for new programs, services, or models? (If so, probe for how these champions influence the adoption process.)

Adoption of FFT and Program Change Experiences

Now we are going to ask you some questions about your particular experience as agency lead in the adoption of FFT. We are especially interested in hearing about the factors that made this an easy or difficult experience for you. We know from talking with OMH that there are opportunities to improve on how the program was rolled out in terms of the preparation and groundwork, technical assistance and technical resources. OMH would like to know about these and other factors that will help them create a more efficient system for rolling out interventions in the future.

12. What was the impetus behind the adoption of FFT? That is, what were the major reasons why your agency adopted FFT? (Probe for internal and external factors; how much they knew about the FFT model and program ahead of time; perceived value of FFT; fit with agency's mission.)

13. What steps did your organization take in adopting it (e.g., informing internal and external stakeholders regarding the rationale, implications, priority, what implementation would look like; developing an implementation and support infrastructure, including staff selection and support processes, roles and responsibilities, resource allocation, communications and tracking systems, policies and procedures, outcome evaluation)?

14. To what extent do you feel FFT has been adopted successfully in your organization? (Probe for stage of adoption; percent of clinicians providing FFT; FFT completion rates; types of client outcomes from FFT.)

15. What, if any, are the major obstacles that you have encountered while trying to adopt FFT? (Probe for both internal and external factors such as support resources, information technology and computer support, financial viability, problems encountered in maintaining fidelity to the FFT model and clinical aspects of FFT, staffing, incentives for staff, staff norms and attitudes toward working with families, methods of motivating staff and consumers, relationships with referral agencies.)

16. What are the major factors that made adoption of FFT easier or more successful? (Probe for both internal and external factors such as support resources, information technology and computer support, financial viability, problems encountered in maintaining fidelity to the FFT model and clinical aspects of FFT, staffing, incentives for staff, staff norms and attitudes toward working with families, methods of motivating staff and consumers, relationships with referral agencies.)

17. What role have you as agency lead played in the adoption? (Probe for team selection process, internal communications about adopting FFT; fit of FFT with current programs.)

18. What role did the supervisory staff play in the adoption? (Probe for interactions with FFT trainers and other local FFT providers; changes needed in skills, time, leadership, commitment and incentives to implement FFT; importance of identifying a champion for FFT.)

19. What role did clinical staff play in the adoption? (Probe for self-selection to participate, interactions with FFT trainers and other local FFT providers; changes needed in skills, time, leadership, commitment and incentives to implement FFT; importance of identifying a champion or opinion leader for FFT; importance of full-time vs. partial commitment.)

20. What role have consumers played in the adoption? (Probe to their response to FFT.)

21. Was OMH involved in promoting or encouraging adoption of FFT? If yes, in what ways? (Probe for suggestions about what things OMH might do in the future to facilitate adoption of FFT, such as free training and implementation coordination.)

22. Out of all of these groups (or others not listed), were there any clear champions (individuals or groups) for FFT? (Probe for how champions can be identified and encouraged.)

23. If you were beginning the process of adopting FFT now, what steps would you take to better prepare for its adoption?

23a. Would you change staffing patterns (# of staff assigned or supervisory caseloads) to adopt FFT? If yes, in what ways?

23b. Would you change your standard procedures, such as intake assessment, referral, case management, step-down from FFT, supervision, etc.? If yes, in what ways?

23c. Would you change your billing, tracking &and/or documentation procedures? If yes, in what ways?

23d. Are there any other things you can think of that would make it easier to sustain FFT over time?

24. If you were to start the FFT adoption process all over again, would you do anything differently?

25. Do you have any feedback that you would like to share with us about these questions? Are there questions we should be asking that would help us better understand the process by which organizations like yours implement evidence-based practices like FFT?

Appendix F: CYPRESS

Children and Young People - Resources, Evaluation and Systems Schedule (CYPRESS) (v1.10.11)

Rating to be made on basis of information indicated in each section
Please place an X in the box that best describes the current level of service/team functioning

Setting..... Date ..... /..... /..... Assessors ..... Total Score ...../100

Service characteristics	1	2	3	4	5
1. Shared model of care Service has a comprehensive and shared view of the model of care provided which is owned by the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear, shared model(s) of care	Different model(s) of care; not shared by all members of staff	Different model(s) of care but shared by some members of staff	Single model of care shared by most members of staff	Single model of care shared by all members of staff
2. Population served Service has explicit criteria describing the population served including nature of presenting problem(s) accepted, exclusions, age, referral sources. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of population served	Statement of population served, but lacks detail or clarity in relation to population or referral sources	Statement of population served, with detailed and clear criteria for most aspects of population served and referral sources	Statement of population served, with detailed and clear criteria for all aspects of population served and referral sources	Statement of population served, with detailed and clear criteria for all aspects of population served and referral sources which is shared with all referrers
3. Care pathway Service has explicit criteria and processes for referral to, progress through and exit from the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of care pathway or procedures	Statement of pathways through care for some but not all elements of the service provided	Statement of pathways through care for most elements of the service provided with limited exceptions	Statement of pathways through care for all elements of the service provided	Statement of pathways through care for all elements of the service provided which is shared with clients and referrers

4. Service capacity Service has explicit criteria on capacity (caseload) of service (at service and individual staff member level) that are known to service managers and staff, and referrers. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of service capacity	Statement of service capacity but which lacks clarity or detail in one or more aspects	Statement which provides clarity and detail on most aspects of service capacity, with limited exceptions	Statement which provides clarity and detail on all aspects of service capacity	Statement which provides clarity and detail on all aspects of service capacity and which is understood by the team and referrers
5. Relationship to other services Service has clear policies that describe the nature of the relationship to other services (e.g. health, mental health, social, educational, criminal justice, voluntary sector) serving the population covered by the team (e.g. joint working, confidentiality, shared care protocols, different populations served). Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No policies and procedures concerning relationship to other services	Policies and procedures concerning relationship to some services with which the team/service works	Policies and procedures concerning relationship to most services with which the team/service works	Policies and procedures concerning relationship to all services with which the team/service works	Policies and procedures concerning relationship to all services with which the team/service works and in which the nature of the relationship(s) is clearly described
6. Service/team staffing The composition and roles of the staff team is consistent with the tasks required of the team. Staff have defined roles which are well understood by all team members. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of staffing or roles in relation to the tasks of the team	Statement in which the numbers and roles of some but not all of the staff are described in relation to the tasks of the team	Statement in which the numbers and roles of most of the staff are described in relation to the tasks of the team	Statement in which the numbers and roles of all of the staff are described in relation to the tasks of the team	Statement in which the numbers and roles of all of the staff are described in relation to the tasks of the team, which is understood by all team members

Team operation	1	2	3	4	5
7. Team meetings Service has a regular programme of team meetings which have a clear function and structure, address clinical and operational issues and are attended by all staff members Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No regular team meetings	Team meetings occur irregularly and have variable staff attendance	Regular team meetings occur with discussion of clinical or operational issues covered which some staff attend consistently	Regular team meetings occur with clinical or operational issues covered which most staff attend consistently	Regular team meetings occur with both clinical and operational issues covered in full and which all staff attend consistently
8. Supervision Service has a comprehensive model of supervision which provides, for all team members, regular structured (individual or group) supervision in the model operated by the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No supervision provided in the team	Limited supervision provided in the team	Supervision provided for most team members but is not regular	Regular supervision provided for most team members	Regular supervision provided for all team members
9. Staff training All staff are trained in delivery of the model(s) of care and all appropriate interventions. New staff have an induction to ensure competence, and existing staff have opportunities for continuing professional development. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No training provided in the team	Limited or irregular training provided in the team with focus only on particular staff groups (e.g. new staff)	Training provided in the team but not comprehensive, limited to specific groups (e.g. direct care staff)	Regular training provided in the team but not fully comprehensive, does not include all staff	Comprehensive and regular training provided in relevant model/interventions for all staff
10.Team communication Team has clear policies and procedures for communicating information about clients and decisions made by the team, with team colleagues, to clients and with other agencies. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No policies and procedures for team communication	Limited policies and procedures for team communication	Policies and procedures for team communication, but focus is primarily on one area (internal or external communication)	Policies and procedures for team communication, which cover most important areas of internal and external communication	Comprehensive policies and procedures for team communication, which cover all important areas of internal and external communication

11. Client outcome monitoring A programme of routine outcome monitoring is in place which provides feedback to individual clients, staff and the team/service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No procedures for routine outcome monitoring	Limited procedures for routine outcome measurement and no feedback to clients	Clear procedures for routine outcome measurement but with limited feedback to clients	Clear procedures for routine outcome measurement with some feedback in most areas to clients	Comprehensive procedures for routine outcome measurement with comprehensive feedback in all areas to clients
12. Quality assurance A quality assurance programme is in place which ensures team involvement in reviewing and contributing to the development and maintenance of service quality, for example, through audit and quality improvement initiatives of the service and evaluating client experience of the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No procedures for quality assurance	Limited procedures for quality assurance	Clear procedures for quality assurance but with only occasional direct team involvement	Clear procedures for quality assurance with direct team involvement in some but not all aspects of the process	Clear procedures for quality assurance with comprehensive team involvement in all aspects of the process

Delivery of interventions	1	2	3	4	5
13. Range of interventions consistent with model A range of individual, group or family interventions are provided by the team which are consistent with the model of care adopted by the team and meet the needs of clients served by the team. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	Few, if any, appropriate interventions provided	Limited range of appropriate interventions provided to some clients	Good range of appropriate interventions available to most clients	Full range of appropriate interventions provided to most clients	Full range of appropriate interventions provided to all clients
14. Assessment The team offers a comprehensive assessment that includes psychological, family, educational and social needs and the resources of the wider system, and that involves the client, the family/carers and relevant agencies. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No or limited assessment of child/young person's needs or the family/care system	Assessment of child/young person's needs but focus limited primarily to one area with little or no involvement of the family/care system	Assessment of child/young person's needs and resources but focus limited primarily to two areas; some involvement of the family/care system	Full assessment of child/young person's needs and resources which involves the client, and the family/care system	Full assessment of child/young person's needs and resources which involves the client, and the and the family/care system, and integrates the assessment across all areas
15. Individualised care Each client has an individualised, collaborative programme of care that addresses their needs and is shared with the client (and others where appropriate) in writing. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No or limited individualised programmes of care which are not developed in collaboration with the client(s)	Limited individualised programmes of care which are only occasionally developed in collaboration with the clients(s)	Most clients have individualised programmes of care, the contents of which are developed in collaboration with clients(s)	All clients have individualised programmes of care, the contents of which are developed in collaboration with most clients(s)	All clients have individualised programmes of care, the contents of which are developed in collaboration with all clients(s)
16. Family/carer involvement Families/carers are involved in assessment, decision making and the provision of care. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No or limited family/carer involvement	Limited family/carer involvement in assessment, decision making or the provision of care	Most clients' families/carers involved in some aspects of assessment, decision making or the provision of care	All clients' families/carers involved in the majority of aspects of assessment, decision making and the provision of care	All clients' families/carers involved in all aspects of assessment, decision making and the provision of care

17. Assertive engagement The team has an assertive approach to the engagement of clients and families/carers, (e.g. a 'no drop out' policy or a stress on overcoming difficulties in engaging with services). Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No assertive approach to the engagement of clients or families/carers	Limited assertive approach to the engagement of clients or families/carers	Assertive approach to the engagement of some clients or families/carers	Assertive approach to the engagement of the majority of clients and families/carers	Assertive approach to the engagement of all clients and families/carers
18. Interventions provided in a range of settings Interventions (individual, group, family) are provided in a range of settings to facilitate access to, uptake of and engagement with services and to enhance the effectiveness of the interventions. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	Provision of all interventions limited to one setting (e.g. health care) only	Provision of some interventions in a limited number of different settings	Provision of most interventions in a number of different settings	Provision of all interventions in a number of different settings	Provision of all interventions in a wide range of different settings as appropriate for all clients
19. Risk and child protection The team has and uses a single integrated comprehensive policy for the identification and management of risk and child protection issues. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No single comprehensive risk and child protection policy in place	A risk and child protection policy is in place which is used and understood by some staff	A single agency comprehensive risk and child protection policy is in place which is used and understood by most staff	A comprehensive integrated risk and child protection policy is in place which is used and understood by most staff	A comprehensive integrated risk and child protection policy is in place which is used and understood by all staff
20. Case management The team has a common approach to the management and coordination of care (i.e. assessment, intervention and follow-up) for all clients. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No common approach to the management of care is in place	A number of approaches to the management of care are in place for clients	A single common approach to care is in place which is used by most staff	A single common comprehensive approach to the management of care is in place which is used for most clients and by all staff	A single common comprehensive approach to the management of care is in place which is used for all clients and by all staff

Appendix G: Letter giving ethical approval

South East Research Ethics Committee
South East Coast Strategic Health Authority
Preston Hall
Aylesford
Kent
ME20 7NJ

Telephone: 01622 713097
Facsimile: 01622 885966

20 May 2009

Professor Peter Fonagy
Freud Memorial Professor of Psychoanalysis and Head of the Research Department of Clinical, Educational and Health Psychology, University College London
University College London
Psychoanalysis Unit
1-19 Torrington Place
UCL
WC1E 7HB

Dear Professor Fonagy

Full title of study: **START (Systemic Therapy for At Risk Teens): A National Randomized Controlled Trial to Evaluate Multisystemic Therapy in the UK Context**
REC reference number: **09/H1102/55**

The Research Ethics Committee reviewed the above application at the meeting held on 13 May 2009.

After the Committee's initial deliberations on your application, yourself and Dr Butler kindly joined the meeting to clarify some issues. Thank you for taking the time to do so. The following issues were clarified during the discussion:

Q **Can you deliver this? It is a very intensive process with many contacts with members of families. Do you have enough resources?**

A There is clinical provision in place within the ten established sites. A government grant of £10million has been awarded to this project. All staff have already been recruited for the ten sites. Seven sites have staff employed by NHS agencies and three have staff employed by local authorities. Collaboration was demonstrated in order to gain the funding. All systems necessary have already been developed. The study will be monitored very carefully to ensure intervention is properly delivered.

Q **Has risk assessment taken into account that you may not be able to undertake the project exactly as per the proposal?**

A One of the outcome variables is to expect site-specific differences and this should be the guiding principle of any government national roll-out.

Q **In the power calculation you have allowed for differences in sites, but vulnerable young people come from different sources. There will be a multifaceted group receiving the intervention. Has this been taken into account?**

A The power calculation is based on the success rates of the USA and Norway studies primarily recruited from offender centres. There are no figures to inform the power calculation, although most young people will probably be the same regardless of the service they come from. They will all be diligent rejecters from an early age and we have factored in that they may respond less well. Randomisation has been agreed by the funders.

The Committee were very impressed with the thought that had gone into the study, and the helpful attendance of two of the most senior members of the team; and noted that it was very helpful to have received comments from the study reviewers.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission at NHS sites ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
The Revised Conflict Tactics Scales (CTS2)		
Beliefs and Attitudes Scale		
The Development and Well-Being Assessment - Parent Interview		
The Development and Well-Being Assessment - Interview with 11-17 year olds		
Insurance Certificate		01 July 2008
Participant Consent Form: Young Person	1.1	07 April 2009
Participant Information Sheet: Parent or Carer	1.1	07 April 2009
Participant Information Sheet: Young People aged 15-17	1.1	07 April 2009
Participant Information Sheet: Young People aged 11-14	1.1	07 April 2009
Questionnaire: Strengths and Difficulties Questionnaire		
Questionnaire: The University of New Orleans Alabama Parenting Questionnaire (APQ)		
Questionnaire: Short Mood and Feelings Questionnaire		
Questionnaire: The General Health Questionnaire		
Questionnaire: Young Person's Questionnaire Booklet		
Peer Review		
Letter from Sponsor		04 April 2009

Covering Letter		08 April 2009
Protocol	1.0	30 March 2009
Investigator CV	Professor Peter Fonagy	
Application		07 April 2009
Connors' Teacher Rating Scale - Revised (S)		
ICU (Youth Version)		
ICU (Parent Version)		
LEE scale		
The McMaster Family Assesment Device		
WASI Record Form		
The Child Attachment Interview (CAI) Protocol		
Participant Consent Form: Parent/Carer	1.1	07 April 2009
Participant Consent Form: Optional Additional Qualitative Study - Parent/Carer	1.1	07 April 2009
Participant Information Sheet: Optional Additional Qualitative Study Information for Parents	1.1	07 April 2009
Development and Well-being Assessment (Teacher Version)		

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Professor Katona and Dr Bhiman both declared a non-specific, non-personal interest in the study. Members agreed that Professor Katona and Dr Bhiman could remain in the meeting and contribute to the review of the study.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document “After ethical review – guidance for researchers” gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

09/H1102/55	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

Dr L. Alan Ruben
Chair

Email: nicki.watts@nhs.net

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
"After ethical review – guidance for researchers"

Copy to: *Dr O Avwenagha*

South East Research Ethics Committee

Attendance at Committee meeting on 13 May 2009

Committee Members:

Name	*Profession*	*Present*	*Notes*
Dr Dipti Amin	Physician	Yes	
Dr Ashok Bhiman	Consultant Psychiatrist	Yes	
Dr Bob Brecher	Reader in Moral Philosophy	Yes	
Professor David Caplin	Physicist	Yes	
Professor David Croisdale-Appleby	Professor in Medical Research and Medical Education	Yes	
Professor John Eastwood	Consultant Renal Physician	Yes	
Dr Alan Fishtal	GP	Yes	
Dr Anne Gallagher	Senior Research Fellow (Nurse Member)	No	
Mr Guy Gardener	Retired Assistant Chief Constable	Yes	
Dr Ray Godfrey	Educational Statistician	No	
Mrs Vera Hughes	Training Consultant	Yes	
Dr Anton Joseph	Consultant Radiologist	Yes	
Professor Cornelius Katona	Academic Psychiatrist	Yes	
Dr Robin MacKenzie	Director Medical Law & Ethics	No	
Professor Liz Meerabeau	University Professor (Nurse Member)	No	
Dr L. Alan Ruben	GP	Yes	
Mr Roy Sinclair	Pharmacist	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Miss Nicki Watts	Co-ordinator
Mr Ron Driver	University Lecturer (Observing)

Appendix H: CYPRESS manual

CYPRESS
Children and Young People - Resources, Evaluation and Systems Schedule

Administration Manual

1. Introduction

CYPRESS is designed to describe the structure, operations and interventions provided by services to children and young people with complex needs. It is designed for use as a measure of fidelity in the delivery of services for children and young people with complex needs. Its primary use will be as a research tool but services may wish to use it as part of a review of their functioning. The scale addresses three key aspects of service provision. They are:

- Service characteristics
- Team operation
- Delivery of interventions

Service characteristics – this section is concerned with the overall service ethos and philosophy, the population served, the pathways into care, service capacity, and the relationship with other services and service staffing.

Team operation – this section covers the programme of staff meetings, systems for supervision and staff training, and communications systems within and external to the team. In addition there are sections on client outcome monitoring, and quality assurance.

Delivery of interventions - this section is concerned with the range of interventions provided by the service, the assessment procedures, the delivery of individualised care, family and carer involvement, assertive engagement, the provision of interventions across a range of settings, child protection, and case coordination.

The scale has 20 items, each scored from 1 to 5 giving a maximum total score of 100; for each section scores of 30 (for service characteristics), 30 (for team operation) and 40 (for delivery of interventions), are possible.

2. Administration

The scale is designed to be administered and rated by two individuals. This is because of the complex nature of the services under review and because the process of data collection is likely to be more reliable and valid if undertaken by two people rather than a single person. It is expected that at least one of the individuals involved in the data collection be experienced in the delivery of child and adolescent mental health services.

Data collection requires a number of sources; these include direct interviews with team leaders/supervisors and staff providing the services; and a review of written material, including operational policies, audit reports, activity reports, and information on systems. In addition there maybe contact with external agencies working with the services where required.

As can be seen from review of the scale (Appendix 1) it covers a broad range of activity and structures of a service and requires detailed information from a range of sources for its

completion. The collection of the data and its evaluation involve a number of stages, they are summarised below:-

i) Initial contact with service

A letter explaining the purpose of the scale should be sent to the service. This will normally follow a discussion with the service (including senior managers), at which the agreement to use the scale has been confirmed. (Note this where staff external to the service for research or service evaluation purposes use the scale but it may not be needed where the scale is used for an internal review of the service).

The initial letter should be accompanied by the pre-interview information form (see Appendix X) to request from the service the following:

a) Copies of all relevant operating policies and procedures in the areas addressed by the scale

b) Information on areas such as number of staff, team meetings and organisational arrangements

c) A request for interviews with relevant staff of the services should be made including:
   - Team leader level
   - Staff providing the service

It is anticipated that this contact with the services normally be made at least 4 weeks in advance of the data being required to give the services time to collect data and organise personnel for the interviews.

ii) Initial service rating

Prior to the assessors meeting the service the assessors should normally make an initial rating of the service on the basis of the written information provided. The rating will be made independently by both raters and where discrepancies arise they will be resolved by discussion; where this is not possible a decision will not made until more information is available. Ratings (see Appendix 3 for further information on making the rating) will be provisional and will be adjusted in light of the interviews with staff and any further written information that is made available. A key purpose of the provisional rating will be to guide the questions for staff and any subsequent information requests to the service.

iii) Meeting with staff of the service

Wherever possible it is suggested that the meeting with the service takes place on a single day. A minimum one-hour meeting (but allowing of up to 90 minutes) should be arranged with the following:

- Team Leader(s) of the service
- Staff of the service

Additional time may also be required to access and review new data/information that the services have provided.

The purpose of these interviews/data reviews is to consider all 20 items on the scale with relevant staff of the service. The precise nature and structure of the questions will be influenced by the information obtained from the written material already provided. In some cases the questions may just be points of confirmation and/or clarification. In other cases, in the absence of adequate written information, further basic information may be sought.

iv) *Final service rating*

As the interviews are undertaken and any additional information collected, the two raters should independently rate the service on each item. This may confirm or adjust the initial rating. As with the initial rating the two assessors should review and rate each item individually and agree a consensus rating. Where a consensus is not possible assessors may need to seek further information or clarification from the service in question, followed by seeking advice from a senior colleague to resolve the difficulty if necessary. Where a rating cannot be made a zero is scored.

v) *Feedback*

All services should be offered feedback on the outcome of the review. This should generally be written in a positive and encouraging manner highlighting both the strengths and weaknesses of the service. (The timing of the feedback will be influenced by the purpose for which the rating was undertaken; for example in a research study it may be necessary to delay feedback of the data until completion of the study).

Appendix 1

Children and Young People - Resources, Evaluation and Systems Schedule (CYPRESS) (v1.10.11)

Rating to be made on basis of information indicated in each section

Please place an X in the box that best describes the current level of service/team functioning

Setting..... Date ..... /..... /..... Assessors ..... Total Score ...../100

Service characteristics	1	2	3	4	5
1. Shared model of care Service has a comprehensive and shared view of the model of care provided which is owned by the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear, shared model(s) of care	Different model(s) of care; not shared by all members of staff	Different model(s) of care but shared by some members of staff	Single model of care shared by most members of staff	Single model of care shared by all members of staff
2. Population served Service has explicit criteria describing the population served including nature of presenting problem(s) accepted, exclusions, age, referral sources. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of population served	Statement of population served, but lacks detail or clarity in relation to population or referral sources	Statement of population served, with detailed and clear criteria for most aspects of population served and referral sources	Statement of population served, with detailed and clear criteria for all aspects of population served and referral sources	Statement of population served, with detailed and clear criteria for all aspects of population served and referral sources which is shared with all referrers
3. Care pathway Service has explicit criteria and processes for referral to, progress through and exit from the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of care pathway or procedures	Statement of pathways through care for some but not all elements of the service provided	Statement of pathways through care for most elements of the service provided with limited exceptions	Statement of pathways through care for all elements of the service provided	Statement of pathways through care for all elements of the service provided which is shared with clients and referrers

4. Service capacity Service has explicit criteria on capacity (caseload) of service (at service and individual staff member level) that are known to service managers and staff, and referrers. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of service capacity	Statement of service capacity but which lacks clarity or detail in one or more aspects	Statement which provides clarity and detail on most aspects of service capacity, with limited exceptions	Statement which provides clarity and detail on all aspects of service capacity	Statement which provides clarity and detail on all aspects of service capacity and which is understood by the team and referrers
5. Relationship to other services Service has clear policies that describe the nature of the relationship to other services (e.g. health, mental health, social, educational, criminal justice, voluntary sector) serving the population covered by the team (e.g. joint working, confidentiality, shared care protocols, different populations served). Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No policies and procedures concerning relationship to other services	Policies and procedures concerning relationship to some services with which the team/service works	Policies and procedures concerning relationship to most services with which the team/service works	Policies and procedures concerning relationship to all services with which the team/service works	Policies and procedures concerning relationship to all services with which the team/service works and in which the nature of the relationship(s) is clearly described
6. Service/team staffing The composition and roles of the staff team is consistent with the tasks required of the team. Staff have defined roles which are well understood by all team members. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No clear statement of staffing or roles in relation to the tasks of the team	Statement in which the numbers and roles of some but not all of the staff are described in relation to the tasks of the team	Statement in which the numbers and roles of most of the staff are described in relation to the tasks of the team	Statement in which the numbers and roles of all of the staff are described in relation to the tasks of the team	Statement in which the numbers and roles of all of the staff are described in relation to the tasks of the team, which is understood by all team members

Team operation	1	2	3	4	5
7. Team meetings Service has a regular programme of team meetings which have a clear function and structure, address clinical and operational issues and are attended by all staff members. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No regular team meetings	Team meetings occur irregularly and have variable staff attendance	Regular team meetings occur with discussion of clinical or operational issues covered which some staff attend consistently	Regular team meetings occur with clinical or operational issues covered which most staff attend consistently	Regular team meetings occur with both clinical and operational issues covered in full and which all staff attend consistently
8. Supervision Service has a comprehensive model of supervision which provides, for all team members, regular structured (individual or group) supervision in the model operated by the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No supervision provided in the team	Limited supervision provided in the team	Supervision provided for most team members but is not regular	Regular supervision provided for most team members	Regular supervision provided for all team members
9. Staff training All staff are trained in delivery of the model(s) of care and all appropriate interventions. New staff have an induction to ensure competence, and existing staff have opportunities for continuing professional development. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No training provided in the team	Limited or irregular training provided in the team with focus only on particular staff groups (e.g. new staff)	Training provided in the team but not comprehensive, limited to specific groups (e.g. direct care staff)	Regular training provided in the team but not fully comprehensive, does not include all staff	Comprehensive and regular training provided in relevant model/interventions for all staff
10.Team communication Team has clear policies and procedures for communicating information about clients and decisions made by the team, with team colleagues, to clients and with other agencies. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No policies and procedures for team communication	Limited policies and procedures for team communication	Policies and procedures for team communication, but focus is primarily on one area (internal or external communication)	Policies and procedures for team communication, which cover most important areas of internal and external communication	Comprehensive policies and procedures for team communication, which cover all important areas of internal and external communication

11. Client outcome monitoring A programme of routine outcome monitoring is in place which provides feedback to individual clients, staff and the team/service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No procedures for routine outcome monitoring	Limited procedures for routine outcome measurement and no feedback to clients	Clear procedures for routine outcome measurement but with limited feedback to clients	Clear procedures for routine outcome measurement with some feedback in most areas to clients	Comprehensive procedures for routine outcome measurement with comprehensive feedback in all areas to clients
12. Quality assurance A quality assurance programme is in place which ensures team involvement in reviewing and contributing to the development and maintenance of service quality, for example, through audit and quality improvement initiatives of the service and evaluating client experience of the service. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No procedures for quality assurance	Limited procedures for quality assurance	Clear procedures for quality assurance but with only occasional direct team involvement	Clear procedures for quality assurance with direct team involvement in some but not all aspects of the process	Clear procedures for quality assurance with comprehensive team involvement in all aspects of the process

Delivery of interventions	1	2	3	4	5
13. Range of interventions consistent with model A range of individual, group or family interventions are provided by the team which are consistent with the model of care adopted by the team and meet the needs of clients served by the team. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	Few, if any, appropriate interventions provided	Limited range of appropriate interventions provided to some clients	Good range of appropriate interventions available to most clients	Full range of appropriate interventions provided to most clients	Full range of appropriate interventions provided to all clients
14. Assessment The team offers a comprehensive assessment that includes psychological, family, educational and social needs and the resources of the wider system, and that involves the client, the family/carers and relevant agencies. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No or limited assessment of child/young person's needs or the family/care system	Assessment of child/young person's needs but focus limited primarily to one area with little or no involvement of the family/care system	Assessment of child/young person's needs and resources but focus limited primarily to two areas; some involvement of the family/care system	Full assessment of child/young person's needs and resources which involves the client, and the family/care system	Full assessment of child/young person's needs and resources which involves the client, and the and the family/care system, and integrates the assessment across all areas
15. Individualised care Each client has an individualised, collaborative programme of care that addresses their needs and is shared with the client (and others where appropriate) in writing. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No or limited individualised programmes of care which are not developed in collaboration with the client(s)	Limited individualised programmes of care which are only occasionally developed in collaboration with the clients(s)	Most clients have individualised programmes of care, the contents of which are developed in collaboration with clients(s)	All clients have individualised programmes of care, the contents of which are developed in collaboration with most clients(s)	All clients have individualised programmes of care, the contents of which are developed in collaboration with all clients(s)
16. Family/carer involvement Families/carers are involved in assessment, decision making and the provision of care. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No or limited family/carer involvement	Limited family/carer involvement in assessment, decision making or the provision of care	Most clients' families/carers involved in some aspects of assessment, decision making or the provision of care	All clients' families/carers involved in the majority of aspects of assessment, decision making and the provision of care	All clients' families/carers involved in all aspects of assessment, decision making and the provision of care

17. Assertive engagement The team has an assertive approach to the engagement of clients and families/carers, (e.g. a 'no drop out' policy or a stress on overcoming difficulties in engaging with services). Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No assertive approach to the engagement of clients or families/carers	Limited assertive approach to the engagement of clients or families/carers	Assertive approach to the engagement of some clients or families/carers	Assertive approach to the engagement of the majority of clients and families/carers	Assertive approach to the engagement of all clients and families/carers
18. Interventions provided in a range of settings Interventions (individual, group, family) are provided in a range of settings to facilitate access to, uptake of and engagement with services and to enhance the effectiveness of the interventions. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	Provision of all interventions limited to one setting (e.g. health care) only	Provision of some interventions in a limited number of different settings	Provision of most interventions in a number of different settings	Provision of all interventions in a number of different settings	Provision of all interventions in a wide range of different settings as appropriate for all clients
19. Risk and child protection The team has and uses a single integrated comprehensive policy for the identification and management of risk and child protection issues. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No single comprehensive risk and child protection policy in place	A risk and child protection policy is in place which is used and understood by some staff	A single agency comprehensive risk and child protection policy is in place which is used and understood by most staff	A comprehensive integrated risk and child protection policy is in place which is used and understood by most staff	A comprehensive integrated risk and child protection policy is in place which is used and understood by all staff
20. Case management The team has a common approach to the management and coordination of care (i.e. assessment, intervention and follow-up) for all clients. Evidence from interviews with/reviews of: Team leader(s) Team members Written materials	No common approach to the management of care is in place	A number of approaches to the management of care are in place for clients	A single common approach to care is in place which is used by most staff	A single common comprehensive approach to the management of care is in place which is used for most clients and by all staff	A single common comprehensive approach to the management of care is in place which is used for all clients and by all staff

Appendix 2

CYPRESS Checklist

Completing the checklist

This checklist is designed to help in the completion of the CYPRESS. It is organised into three areas, which relate to the structure of the scale and the way in which the data will be collected. This data will be reviewed to help complete the scale and will be supplemented by questions for managers and staff of the service. The main areas to be covered are set out in column 1 and possible sources of data are given as examples in column 2. Please supply any information (policies, reports, standard data reports), which would allow for the best assessment of your service. Please list the information you supply in column 3. This may result in some duplication of items listed in column three but this is helpful in completing the assessment. Please attach a copy of the relevant documentation to the checklist. (The checklist and documentation may be submitted electronically or in hard copy).

Service/team …… Date submitted ……/……/……

Area	Possible sources of information	Documents supplied by your services (Please attach to this checklist)
Service characteristics including: • Service ethos and comprehensiveness • Population served • Care pathway • Service capacity • Relationship to other services • Service staffing and monitoring	• Operational policies • Information about the service • Reports (internal and external on the service) • Service evaluations and audit reports • Referral/assessment systems • Staffing structures (including full and vacant post) • Job descriptions • Care protocols • Information sharing protocols	1. 2. 3. 4. 5. 6.
Team operation • Team meetings • Shared model of care • Staff supervision and training • Team communication • Client outcome monitoring • Quality assurance • Feedback systems	• Operational policies • Team policies, meeting structures and minutes • Reports (internal and external on the service) • Supervision policy • Service evaluations and quality reports • Feedback mechanisms and reports • Staff training policy and programme of training • Outcome monitoring tools and reports	1. 2. 3. 4. 5. 6.

Delivery of interventions including: • Range of interventions • Assessment • Individualised care • Family/carer involvement • Assertive engagement • Interventions across a range of sites/settings • Risk and child protection • Case coordination • Shared decision making	• Description of interventions provide • Structure of assessment (and assessment tools used) • Policy for and any individual care planning tools • Policy and procedure for assertive engagement • Lettings in which service is provided • Feedback mechanisms and reports • Child protection policy and procedure(s) • Case coordination policy and relevant materials • Shared decision making policy and materials	1. 2. 3. 4. 5. 6.

Appendix 3

CYPRESS Pre-interview Information Form

Prior to meeting with your team we would like to request some information that will help us with the process and hopefully reduce the duration of interviews and help us to gain a more complete understanding of the way the team/service works. There are two sections: the first section covers written documentation which would be very helpful to see before the interviews, and the second section seeks some basic information about the service which will help with conducting the interviews.

If you have any questions about completing this form please contact X.

Team/service: ……

Team/service manager: ……

Section 1 – DOCUMENTATION

Please provide copies of the following documents where available:

- Operational policy
- Referral and service protocol(s)
- Annual report
- Audit reports

Also the following, where not included in operational policy:

- Clinical risk management and safeguarding policy
- Supervision policy
- Training policy

Section 2 - INFORMATION ABOUT THE TEAM/SERVICE

Team development and operation	
For how long (months/years) has the team been in operation in its current form?	
What type of organisation does the team sit within (e.g. NHS, local authority, third sector, joint arrangements)?	
Is the team a 'stand-alone team' or is it integrated with other services?	
What is/are the source(s) of funding to the team?	
Staffing	
Number of clinical staff (WTE)	
Number of administrative staff (WTE)	
Please list the different professional and support roles in the team and number of whole time equivalents	
Number of new starters in the last 24 months (WTE)	
Number of staff who have left the team in the last 24 months (WTE)	

Caseload	
Current number of clients on team caseload	
Maximum number of clients on team caseload	
Maximum individual clinician caseload number	
Current average individual clinician caseload number	
Referrals	
Average wait between referral and initial appointment	
Supervision	
Please describe the type and frequency of supervision (e.g. individual, group, case management, clinical supervision, case discussion, video review, etc)	
Team meetings	
Please describe frequency and duration of routine team meetings	

Please send any documents and the completed CYPRESS pre-interview information form either electronically to X or by post to X. Many thanks.

Appendix 4

Achieving reliable ratings

Where ratings are based on a percentage of staff/clients who have, for example, undertaken a task or been offered an intervention, the following act as a guide to rating the services.

All	= <95%
Most	= 70% - 95%
Majority	= 50% - 69%
Some	= 30% - 49%

SERVICE CHARACTERISTICS

1. Shared model of care
The service has a clear, comprehensive shared view of the model of care provided that is owned and understood by the service and all team members. A statement of this model of care should be the operationalisation of the service ethos. Where models of care are known but not necessarily shared by staff members, potential conflicts arising from this will result in a lower score. To score highly on this item the team should be able to articulate consistently the shared model if care adopted by the service.

2. Population served
The service has clear and explicit written criteria that describe the population served. Typically this will involve the specification of the presenting problems treated by the service, any exclusions (and the justification for them), the age range and developmental needs of individuals served, and potential referral sources. This should be available to external services, and children and their families. These criteria should be understood by the staff of the service. There should be little or no disagreement on that the criteria as applied in practice. It is also important in this to rate how well implemented the strategy is.

3. Care pathway
This refers to the processes by which a referral is made, its progress through the service and how an exit from the service is achieved. Pathways should be clear, specifying criteria for entry into, movement through, and exit from the service. This should be specified in a way that is clear to all staff, referrers, clients and carers. This should be in written form and available to stakeholders.

4. Service capacity
The service has an established capacity expressed in terms of the overall maximum caseload of the service and the individual clinician maximum caseload (including part time equivalents). This is not an attempt to assess the demand on a service but rather whether the service has a clear statement of its capacity that can be understood by those both working in and referring to the service.

5. Relationship to other services
This concerns the policies which describe the nature of the relationship of the service to other health, mental health, social, educational, criminal justice and other services working with the target population. This may be expressed in protocols for, for example, joint working, confidentiality and shared care. To score well on this item the protocol should be clearly expressed in written form. It is important that the relationships are understood by all members of the team.

6. Service/team staffing
The objective here is to establish whether the composition and roles of the staff team is consistent with the tasks required of the team, and the degree to which roles are understood by all team members. A full and clear statement of the staffing, with defined competencies, where roles are understood by all members of the service, would be required to score well.

TEAM OPERATION

7. Team meetings
This item assesses whether the service has a regular programme of team meetings that are concerned with clinical and operational issues addressed, and which are attended by all staff members and minuted. The absence of team meetings would lead to a score of zero. To score well on this measure the team meetings need to deal with clinical and operation issues, with good staff attendance and good records of the meetings.

8. S*upervision*
The service has a comprehensive system for the provision of regular supervision, which provides for all team members. To score well on this item this should be on at least a fortnightly basis and provided for all team members working at all levels.

9. Staff training
Staff training should be delivered in the interventions(s) relevant to model of care. There should also be processes in place to ensure that all staff are competent, have continual professional development and refresher training. The absence of any such training will result in a low score. Comprehensive training will be characterised as training in which elements are provided to all staff members, for example induction and updates on team policy and developments in the models of care, but also allows for specific training in competencies provided by the team.

10. Team communication
The team has clear and explicit policies in place to communicate about clients and decisions made by the team, with other team members, clients and carers and agencies external to the team. These policies should be clearly written and focused on both internal and external communication. To score well on this measure the communication policies should be owned and understood by all staff members.

11. Client outcome monitoring
A programme of routine outcome monitoring is in place that provides feedback to clients, staff and referrers. Low scores will be obtained where there are no processes in place for routine outcome monitoring (defined as sessional or near sessional reporting of individual outcomes). In order to score well on this measure there need be not only routine outcome measurement in place but also clear procedures for feedback at all levels. This feedback should be available to individual clients and carers, the team and the wider service.

12. Quality assurance
Quality assurance is defined as a process for improving the overall quality of the service, for example through audit or other quality initiatives. For example, evaluating the staff composition against agreed targets or required competencies, the number of patients as against agreed targets, the outcomes as against agreed targets and client experience of the service. Processes should be in place that allows for this data to be collected and discussed directly with the team, concerning all or some aspects of the programme.

DELIVERY OF INTERVENTIONS

13. Range of Interventions consistent with a model.
This might include psychological, psychosocial, pharmacological, educational and vocational interventions. The range of interventions should be identified within the operational policy. In addition the extent to which interventions are actually provided and made available to all clients should be assessed.

14. Assessment
The team offers a comprehensive assessment appropriate to the client group and includes psychological, educational and social needs, and the resources of the wider system available to the client with other agencies actively involved in the process. A low score would be given when an assessment involves only the individual client. To score highly on this item there needs to be a full and comprehensive assessment covering all areas identified above and which involves all relevant parties.

15. Individualised care
Each client has individualised care package that is developed in collaboration with the client. To score highly a team must be able to demonstrate, through examples, the way in which their work is both individualised and collaborative.

16. Family and carers involvement
Families and carers are involved in the assessment, provision of care and decision making throughout the course of intervention. To score highly on this item respondents must be able to provide examples of the ways in which families and carers are involved.

17. Assertive engagement
Individuals with complex needs often lose contact with services. The service should have in place a policy for assertive engagement which endeavours to engage clients who might have disengaged with other services in the past. This might be exemplified by a 'no drop out' policy or a flexible engagement approach. For the service to score well such an approach will be taken for all clients and not adopted for only a minority of clients.

18. Interventions provided in a range of settings
This item assesses the degree to which services are available in a range of settings (e.g. different venues in health, social care and community settings) and which are determined by the individual clients and families. It makes a distinction between a site and setting. Variation of site but not in type of setting would result in a lower score.

19. Risk and child protection
The service has a single comprehensive policy for the management of child protection issues. Crucial are the presence of a comprehensive shared model across the whole service as opposed to the existence of a policy relating to a particular organisation which may be differentially operating within the team, or the absence of a policy.

20. Case management
This describes an approach to the management and coordination of care that integrates the assessment of and provision of care for all clients in a standardized manner. A common agreed policy for case coordination would be required for a high score which is in place for all clients and understood and implemented by all staff.

Appendix I: CYPRESS pre-interview information form

CYPRESS Pre-interview Information Form

Prior to meeting with your team we would like to request some information that will help us with the process and hopefully reduce the duration of interviews and help us to gain a more complete understanding of the way the team/service works. There are two sections: the first section covers written documentation which would be very helpful to see before the interviews, and the second section seeks some basic information about the service which will help with conducting the interviews.

If you have any questions about completing this form please contact the START research team – Cressida Gaffney (cressida.gaffney.09@ucl.ac.uk) or Stephen Butler (stephen.butler@ucl.ac.uk or 020 7679 5985).

Service/team: ……

Service/team manager: ……

Section 1 - DOCUMENTATION

Please provide copies of the following documents where available:

- Operational policy
- Referral and service protocol(s)
- Annual report
- Audit reports

Also the following, where not included in operational policy:

- Clinical risk management and safeguarding policy
- Supervision policy
- Training policy

Section 2 - INFORMATION ABOUT THE TEAM

Team development and operation	
For how long (months/years) has the team been in operation in its current form?	
What type of organisation does the team sit within (e.g. NHS, local authority, third sector, joint arrangements)?	
Is the team a 'stand-alone team' or is it integrated with other services?	
What is/are the source(s) of funding to the team?	

Staffing	
Number of clinical staff (WTE)	
Number of administrative staff (WTE)	
Please list the different professional and support roles in the team and number of whole time equivalents	
Number of new starters in the last 24 months (WTE)	
Number of staff who have left the team in the last 24 months (WTE)	
Caseload	
Current number of clients on team caseload	
Maximum number of clients on team caseload	
Maximum individual clinician caseload number	
Current average individual clinician caseload number	
Referrals	
Average wait between referral and initial appointment	
Supervision	
Please describe the type and frequency of supervision (e.g. individual, group, case management, clinical supervision, case discussion, video review, etc)	
Team meetings	
Please describe frequency and duration of routine team meetings	

Please send any documents and the completed CYPRESS pre-interview information form either electronically to cressida.gaffney.09@ucl.ac.uk or by post to Cressida Gaffney, Research Department of Clinical, Educational and Health Psychology, University College London, Gower Street, London, WC1E 6BT. Many thanks.

Appendix J: CYPRESS participant information sheet

CYPRESS (The Children and Young People - Resources, Evaluation and Systems Schedule)

Thank you for agreeing to take part in this exercise which constitutes part of the START (Systemic Therapy For At Risk Teens) research evaluation of a form of intervention for young people and their families who are experiencing difficulties at home, at school and sometimes with the law. As you will be aware, the study is one the largest of its kind, evaluating the effectiveness of clinical services across 9 sites in the UK, representing a collaboration between University College London, the University of Cambridge and the University of Leeds, funded by the Department of Health and the Department of Education.

This measure, CYPRESS, is aimed at characterising services offering complex interventions for children and young people with conduct disorder and other behavioural problems. CYPRESS aims to understand to what degree particular characteristics are intrinsic to such services, and what relationships exist between these particular service characteristics and clinical outcomes.

CYPRESS is divided into three overarching areas in its characterisation of services: service characteristics, team operation, and interventions. CYRESS is not a measure of individual staff performance or specific client outcomes, rather it aims to characterise systemic features of a service taking account of information from staff working at different levels and from organisational documentation. Feedback will be provided to participating teams.

Many thanks for your participation. Should you have any questions please contact the CYPRESS team on cressida.gaffney.09@ucl.ac.uk .