Received: 5 December 2017 | Revised: 19 March 2018 | Accepted: 23 June 2018

DOI: 10.1002/capr.12184

ORIGINAL ARTICLE

WILEY

Measuring the effectiveness of individual therapy on the well-being of children and young people who have experienced abusive relationships, particularly domestic violence: A case study

Sarah Perry[1] | **Ian Frampton**[2]

[1]CLEAR, Truro, Cornwall, UK

[2]School of Psychology, Washington-Singer Laboratory, University of Exeter, Exeter, UK

Correspondence
Sarah Perry, CLEAR,
Cornwall, UK.
Email: sarahperrycornwall@gmail.com

Funding information
CLEAR

Abstract

Background: The evidence base regarding effective interventions for children and young people who experience abusive relationships, particularly domestic violence, is extremely limited.

Aims: To evaluate the effectiveness of individual therapy for children and young people who have experienced abusive relationships; to address the challenges in carrying out such research.

Methodology: A nonexperimental service evaluation repeated measures design with the administration of Routine Outcome Measures (ROMs) that assessed, skills, functioning and symptoms and experience of service. In total, 735 participants aged between 4 and 17 years were referred to the service. More than three quarters of those referred, attended at least three sessions and completed therapy.

Results: The majority of those referred to the service presented with high needs. Mean scores on each of the ROMs improved and these changes were statistically significant. Based on clinical threshold scores, there was a deterioration for a small proportion of the sample (10% or less), improvements for between 28.5 and 49.4% and no change for the remainder. Statistical effect sizes, as measured by the ROMs, were generally in the medium range; satisfaction with the service was high.

Implications: Gathering evidence to demonstrate the effectiveness, or not, of therapy is fraught with difficulties. In the research world, evaluation studies of "real world" interventions may be regarded as inferior to gold standard randomised controlled trials. However, such studies may be better at capturing everyday practice and what can be realistically achieved in terms of measuring effectiveness. There are important implications for commissioning and evaluating such interventions.

KEYWORDS
children and young people, domestic abuse, evaluation, individual therapy

1 | INTRODUCTION

Domestic violence and abuse and children's welfare.

Domestic violence and abuse (DVA) between parents is the most frequently reported form of trauma for children and young people in the UK, with nearly one quarter (24.8%) of those aged 18–24 reporting that they experienced domestic violence and abuse during their childhood (Meltzer, Doos, Vostanis, Ford & Goodman, 2009). The most recent UK Government Department of Education report on the characteristics of Children in Need identified domestic violence as the most frequently reported risk factor identified in needs assessments (DfE, 2016). Out of 448,200 assessments undertaken in 2015/16 in England, 222,200 (49.6%) children were identified at risk of domestic violence.

Domestic violence and abuse is one of a number of specific adverse childhood experiences (referred to as ACEs) that are known to affect many aspects of children's development (Holt, Buckley & Whelan, 2008; Stanley, 2011) and mental health (Boeckel, Wagner & Grassi-Oliveira, 2015; Kitzmann, Gaylord, Holt & Kenny, 2003) and result in multiple health risk factors in adulthood (Felitti et al., 1998). Indeed, ACEs have been referred to as the "hidden epidemic" and currently, one of the most significant public health issues of our time (Van der Kolk, 2015).

Unfortunately, influential guidelines in the UK on DVA and child abuse often marginalise the impact of DVA has on children directly. For example, public health guidelines on DVA state that the term DVA refers to:

> any incident or pattern of incidents of controlling, coercive or threatening behaviour, violence or abuse between those aged 16 or over who are, or have been, intimate partners or are family members.
>
> (NICE, 2014)

This definition focuses primarily on abuse and violence within adult relationships. Consequently, less attention is given to the effects of DVA on children and young people. Indeed, only 2 of the 17 recommendations from the report relate specifically to interventions for children. NICE guidance on DVA confusingly re-iterates that, "Violence and abuse perpetrated on children by adults ('child abuse') is not dealt with in this guidance."

The distinction between child abuse and DVA is also made in NICE guidelines on childhood maltreatment (NICE 2009), which states that maltreatment includes "neglect, physical, sexual and emotional abuse, and fabricated or induced illness" but not "intrafamilial violence." Statutory guidance on inter-agency working to safeguard and promote the welfare of children (DfE, 2015), defines, physical abuse, sexual abuse, child sexual exploitation, neglect and emotional abuse in its glossary. Although the term "emotional abuse" includes "seeing or hearing the ill-treatment of another" DVA is not explicitly referred to. Only the most recent NICE guidance on child abuse and neglect (NICE, 2017) includes witnessing domestic violence.

It appears then, that the DVA literature focuses primarily on adults' experiences and the child abuse literature is not always explicit about the inclusion of DVA. Therefore, children who experience DVA can be overlooked on both accounts. Charities that focus on children's welfare and rights are more forthright on these issues. Thus, the National Society for the Prevention of Cruelty to Children, clearly states on its website "Witnessing domestic abuse is child abuse" (www.nspcc.org.uk). Therefore, in this study exposure to domestic violence is considered to be a type of child abuse.

There has been much campaigning for the recognition of complex trauma or traumatic stress in children and young people who have experienced or been exposed to child maltreatment and domestic violence (see The National Child Traumatic Stress Network; www.nctsn.org). Such trauma is different from other types of trauma in that it is: primarily inter-relational; involves chronic traumatic situations rather than acute traumatic events; is ongoing and current rather than historical; and often occurs in the home environment and is hidden, rather than in the wider community and in public. A recent report on child homicides, whereby children have been killed by perpetrators of domestic violence, highlights that there needs to be better research on how the "relationship between domestic abuse and child welfare is understood by the family judiciary and agencies" (Women's Aid, 2017, p5).

The effectiveness of Interventions for children and young people who have experienced DVA.

Four systematic reviews of the literature on interventions for children and young people who have experienced DVA (BCCEWH, 2013; Howarth et al., 2016; NICE, 2014; Rizo, Macy, Ermentrout & Johns, 2011) have reported on the paucity of robust evidence in this area of practice. These reviews indicate that interventions for children and families who have experienced DVA tend to fall into four main types: counselling/psychotherapy, psychoeducational, parenting skills training and advocacy/outreach. What each of these types of intervention may actually consist of can vary tremendously. Thus, interventions are sometimes offered:

- individually or in groups;
- in isolation, in parallel or together with the nonabusive parent;
- in the short-term (i.e., fixed number of sessions) or longer term (as needed);
- delivered by qualified or unqualified professionals;
- based on variety of therapeutic models (e.g., CBT, play therapy, person-centred counselling);
- consisting of one single type of intervention or a package of multiple interventions.

These reviews conclude that a great deal of uncertainty exists as a result of methodological, political and cultural barriers to conducting good quality research. The most recent review concluded:

> In terms of the UK evidence base and service delivery landscape, there was no UK-based trials, few qualitative studies and little widespread service evaluation. Most programmes are group-based psychoeducational interventions. However, the funding crisis in the DVA sector is significantly undermining programme delivery (pvi) and consequently we do not know if the types of support offered to children in the UK and elsewhere are helpful and well received (acceptable), nor whether or not they represent good value for money.
>
> (NIHR, 2016, pxxv)

Current practices and commissioning arrangements influence what to provide and thus where there is more evidence, but not necessarily what to explore or how to produce robust evidence. In addition, the service characteristics, such as multi-modal treatment approaches and packages of care are much more difficult to fit into gold standard research designs, such as randomised controlled trials (RCTs). Such trials primarily developed within medical models of health, whereby patients with discreet diagnoses are provided with specific treatment regimens.

Despite the serious shortcomings in the research base, two of the four systematic reviews have attempted to synthesise the data in order to make recommendations about further research and practice in the UK (NICE, 2014; NIHR, 2016). These reviews make slightly different recommendations, which demonstrates the significant difficulties around synthesising diverse forms of quantitative and qualitative research (NIHR, 2016).

When recommendations have been made, they have tended to be broad in scope. For example, Recommendation 11 in the NICE (2014) report is about specialist DVA service provision for children and young people, though there is little detail on what "specialist" might refer to:

> Interventions should be timely and should continue over a long enough period to achieve lasting effects. Recognise that long-term interventions are more effective.
>
> (p17)

Howarth et al. (2016) report cited five controlled trials that involved psychotherapeutic interventions, which included psychotherapy, cognitive behavioural therapy and play therapy, with the caveat that psychotherapy need not be delivered by trained psychotherapists. This caveat is concerning on a number of levels (ethical, professional, quality assurance and accountability), and may be related to the fact that training was "not reported" in some published trials. Distinctions are also not made between the mode of delivery (individual or group) and the target of delivery (child, parent or both). There appears to be as much heterogeneity within a type of intervention (for instance, psychotherapeutic) as between different types of interventions (for example, psychotherapeutic vs. psychoeducation). Consequently, making recommendations about a particular type of intervention over another is questionable. Howarth et al. (2016) state their findings should be considered as "hypothesis-generating rather than conclusive."

1.1 | Aims of the study

This study aimed to address some of the gaps in the UK evidence highlighted by recent reviews of the literature. It evaluated the effectiveness of individual therapy for children and young people who have experienced abusive relationships. It also addressed the challenges in carrying out such research. Thus, it will present a case study of how one charity, CLEAR®, in the UK, evaluated the effectiveness of the therapy service it provided to children and young people who have experienced abusive relationships, particularly domestic violence. It is hoped that this focus will contribute to the debate on how we advance our understanding of measuring outcomes and children's experience of accessing services (Wolpert et al., 2016).

2 | METHOD

2.1 | Research design

A nonexperimental repeated measures service evaluation design was used. This involved the administration of Routine Outcome Measures (ROMs) at the beginning and end of therapy to measure the effectiveness of this intervention, and the administration of Experience of Service measures at the end of therapy to capture the acceptability of the service. This research design was influenced by the Children and Young People's Access to Psychological Therapies (CYP IAPT) programme (Wolpert et al., 2016).

2.2 | Participants

CLEAR has provided therapy to children and young people who have experienced abusive relationships since 2008. Children are referred to CLEAR from a variety of agencies and professionals, as well as smaller number of self-referrals. From 2013, CLEAR was commissioned by the Local Authority Domestic and Sexual Violence (DASV) specialty services to provide individual therapy for children and young people aged between 0 and 18 years and up to 25 years for those with additional needs. This contract further defined referral pathways and was available to children and young people known to DASV services. This paper will focus on those referred to CLEAR over a 3-year period (April 2013 to March 2016), when standardised commissioning arrangements introduced more uniformity in the service. It will concentrate on those aged between 4 and 17 years who attended at least three sessions with a therapist. This group of participants all received individual therapy. Children younger than 4 years of age were more likely to receive therapy with their parent.

All these participants experienced DVA or other forms of abusive relationships. It is not possible to provide further information on their or their parents' mental health and background circumstances (for instance, household characteristics, poverty, severity of abuse). Many were referred to the service as a result of their nonabusive parent/carer accessing support and Adult Services identifying children within the family who were exposed to and experienced DVA.

There were exclusion criteria around not working with children where the perpetrator of abuse was still living within the family home. However, therapy was offered to children who remained in contact with the perpetrator, which was common. Sometimes children were referred to CLEAR for other types of relational abuse (e.g., bullying at school), but such referrals were not accepted unless DVA was also a known issue within the home.

2.3 | Intervention

CLEAR provides a selected preventative intervention for children and young people exposed to abusive relationships, as it is recognised these sorts of experiences increase the risk of health and behaviour problems. One of the main principles of CLEAR was listening to young people's voices and enabling them to express themselves through different media (for example, art and music) to help them make sense of their experiences and feel more skilled in managing their thoughts, feelings and behaviours. This work was influenced by a number of theories, in particular those relating to child development (Siegel, 1999), attachment (Bowlby, 1989) and trauma (Van der Kolk, 2015). Therapists were also particularly informed about the impact of domestic violence and abuse on children's well-being through in-house training, awareness of national guidelines and self-directed learning.

Children referred to CLEAR were offered individual therapy which was nondirective and generally time-limited. Individual therapy was delivered by a variety of Psychotherapists, including four Art, four Music, one Dance and Movement, and one Drama Psychotherapist, one Play Therapist, one Gestalt Therapist, four Person-Centred Counsellors and two Clinical Psychologists. Decisions regarding allocation of therapists were largely pragmatic, taking into account geographical area, waiting lists/availability, specific therapy requested and matching therapists' strengths and skills with clients' presenting age and difficulties. All were qualified and registered with their own professional bodies (for example, British Association of Art Therapists) as well as with professional organisations (for instance, Health and Care Professions Council). The charity CLEAR, was also registered with the British Association of Counsellors and Psychotherapists (BACP). All therapists had to have previous experience of working with young people as well as appropriate therapeutic qualifications.

Therapists received regular group supervision within CLEAR provided by qualified psychotherapists and a clinical psychologist. Individual supervision was also provided, as required. High standards of practice were maintained by adhering to professional codes of conduct.

Individual therapy for children who had experienced DASV was commissioned by the local Safety Partnership, who agreed to fund up to 12 sessions of therapy in each case. Sessions were up to one hour in length and were organised on a weekly to monthly basis, dependent on clients' needs and situation. The provision of therapy for other children who had experienced abusive relationships was dependent on grants from a number of charitable organisations (Children in Need, Comic Relief). Again, a similar number of sessions was provided. The number of sessions offered to a child could be extended if a therapist was able to make a strong case for doing so and there were available resources. Children and young people generally attended the first and last session with their nonabusive parent/carer and these sessions focused on assessment and reviewing progress, respectively. The remaining sessions focused on individual therapy.

2.4 | Measures

The NIHR (Howarth et al. 2016) report commented on the heterogeneity of outcome measures used in this area of research, with the recommendation for a consensus in the field about a core outcome data set. CLEAR provided a therapeutic service based on need rather than a particular mental health diagnosis. It focused on improving children's resilience rather than reducing symptoms associated with specific mental health difficulties (although it is often hard to untangle these two aspects of well-being). It was important, therefore, that the outcome measures were broad in scope rather than symptom specific. The selection of measures used to evaluate individual therapy was based on a number of criteria, such as:

1. Evidence of good reliability and validity.
2. Nationally recognised and evaluated.
3. Availability of national norms and thresholds to distinguish between scores within different ranges (such as clinical and nonclinical ranges) and enable comparisons with other studies.
4. Appropriate for use with wide age range (4–18 years).
5. Availability of different versions of the questionnaire for children, young people and parents/carers to complete.
6. Ability to measure specific concepts associated with well-being, such as mental health symptoms, resilience, functioning and satisfaction with the service.

The following routine outcome measures (ROM) were selected and administered as instructed in manuals. Not all the measures were suitable for all children and young people.

1. *Emotional Literacy Checklist* (ELC; Faupel,2003). The ELC is an assessment and intervention measure of a child's skills rather than deficits. It covers five key areas of emotional literacy addressed by the UK Social and Emotional Aspects of Learning (SEAL) curriculum, including empathy, motivation, self-awareness, self-regulation and social skills. Different versions of the ELC can be completed by children (pupil) and young people (student), and their parents and teachers. The measure is suitable for 7- to 16-year olds. Scores range from 25 to 100 with higher scores indicating better functioning. Cut-off scores are provided allowing comparison with a national sample.
2. *Outcome Rating Scale* (*ORS*) and Child ORS (CORS). This measure assesses four areas of life functioning known to change as a result of a therapeutic intervention, including personal well-being, interpersonal relationships, social role and overall well-being (Duncan, Sparks, Miller, Bohanske & Claud, 2006). The CORS is suitable for 6–12-year olds and the ORS for 13+ year-olds. It consists of four 0–10 visual analogue scales, with total scores ranging from 0 to 40 and higher scores representing better life functioning. Cut-off

TABLE 1 Demographic characteristics of children and young people who were referred to the service, engaged with therapy and completed outcome measures at the beginning and end of therapy

	Total referrals *N* = 735[a]				
	Completed therapy (3+ sessions) *N* = 542 (76.9%)	**Disengaged with therapy (<3 sessions)** *N* = 163 (23.1%)	**Difference between those who completed vs. disengaged**	**Completed at least one outcome measure at T1 and T2** *N* = 394	**Difference between those who completed therapy vs. completed ROMs**
Sex					
Female	249 (46.0)	77 (47.5)	NS	182 (46.3)	NS
Male	292 (54.0)	85 (52.5)		211 (53.7)	
Total	541 (100.0)	162(100.0)		393 (100.0)	
Age					
4–10 years	342 (63.1)	93 (57.1)	NS	260 (66.0)	NS
11–17	200 (36.9)	70 (42.9)		134 (34.0)	
Total	542 (100.0)	163 (100.00)		394 (100.0)	
Ethnicity					
White	499 (95.6)	145 (93.5)	NS	364 (96.6)	NS
BME	23 (4.4)	10 (6.5)		13 (3.4)	
Total	522 (100.0)	155 (100.0)		377 (100.0)	
Self-reported disability					
Disabled	62 (11.9)	14 (9.0)	NS	44(11.6)	NS
Non-disabled	459 (88.1)	141 (91.0)		334 (88.4)	
Total	521 (100.0)	155 (100.0)		378 (100)	

[a] Missing data for 30 (4.1%) NS = Not Significant, chi-squared test.

scores are provided allowing comparison with a clinical sample. A Reliable Change Index (RCI), indicating a statistically significant change in individual rating scale score, has been calculated as being greater than five (Campbell & Hemsley, 2009).

3. *Strengths and Difficulties Questionnaire* (SDQ; Goodman, 2001). The SDQ is one of most widely used screening measures of child mental health. It measures emotional and behavioural symptoms, including conduct, emotional, hyperactivity/inattention, peer relationship problems and prosocial behaviour. It also assesses the impact of difficulties on a child's life. The SDQ consists of 25 items. There are different versions that can be completed by young people, parents/carers and teachers. It covers the ages 2–17 years. Total scores range from 0 to 40, with higher scores indicating poorer functioning. There are cut-off scores allowing comparison with national and clinical samples.
4. *Experience of Service Questionnaire* (CHI ESQ; Brown, Ford, Deighton & Wolpert, 2014). This measure was developed to assess service satisfaction in Child and Adolescent Mental Health Services and consists of 12 items and 3 open-ended questions that capture what respondents liked about the service, what needed improving and any other comments There are versions for young people and parents/carers to complete and it covers all ages. Total scores range from 12 to 36, with higher scores representing higher satisfaction.
5. *It's Hard to Put into Words* (IHTPIW; Perry & Carpenter, 2016). This is an in-house measure which does not meet the eligibility criteria, but was developed specifically for the service. It assesses children and young people's views on the benefits of therapy. It consists of 15 items and 3 open-ended questions that capture what respondents thought was good about therapy, what they would like to change about therapy and anything else they would like to say. It is suitable for ages 7 years or more. Total scores range from 0 to 45 with higher scores representing greater benefits.

None of the self-report measures were suitable for young children <6 years of age to complete. Instead parents completed measures to assess changes in their children's well-being.

2.5 | Procedures

Therapists were provided with training on the use of outcome measures, and the measures were administered to clients on iPads at the beginning and end of therapy. This was generally the first and penultimate appointments, respectively.

Children were asked to complete the ELC, ORS/CORS and IHTPIW and parents the ELC, SDQ and CHI-ESQ measures.

2.6 | Ethical considerations

The participation and protection of children and young people in services and research are fundamental human rights which need to be attended to by everyone (UNICEF, 1989). Children living in vulnerable situations where abuse, violence, power inequalities and control have been major issues, require therapists, researchers and services to be particularly sensitive to these aspects of practice (Cater & Øverlien, 2014).

Children and young people are often referred to services by adults, so all therapists ensured they were aware of why they had been referred to the service. Therapists also explained how they worked, confidentiality, how therapy might be set up and a child's readiness for therapy. Participants were informed that they could choose not to complete the outcome measures or stop therapy at any time and could work with a different therapist if they preferred. Participants and their parents/carers were asked to give written consent for the confidential use of outcome data in the evaluation of the service and consequent reports and papers.

2.7 | Statistical analyses

Findings from the ROMs are presented as changes in total scores and movement across predefined cut-offs or bands derived from norms of nationally representative samples. The first approach gives an indication of any change in scores; improvement or deterioration in score. The second gives an indication of the proportion of the sample moving between bands (above, below, average) and how they compare with national norms.

3 | RESULTS

3.1 | Sample characteristics

Between 1 April 2013 and 31 March 2016, 735 participants aged between 4 and 17 years were referred to CLEAR. Engagement with therapy was very good, with just over three quarters of those referred (542, 76.9%) attending at least three sessions and completing therapy (Table 1). Participants received between 3 and 40 sessions of therapy, with an average of 12 sessions. There were no statistically significant differences in demographic characteristics between those who engaged or disengaged from the service.

Response rates to the individual ROMs at the beginning of therapy (T1) and end of therapy (T2) varied. Thus, response rates were good at the beginning of therapy (between 72 and 87%) but dropped at the end of therapy (between 56 and 69%). However, there was at least one complete set of outcome data (T1 and T2) for 394/542 (72.7%) of those children who completed therapy. There were no statistically significant differences in demographic variables between those who completed at least one set of outcome measures and those who did not (Table 1.)

3.2 | Findings from ROMs

Changes in Total Scores for the Emotional Literacy Checklist are shown in Figure 1. Changes in scores were all positive, indicating improvement in emotional literacy, and these changes were all statistically significant for the child ($t = -5.6$, $p < 0.0001$) and parent-reported versions ($t = -12.0$, $p < 0.0001$). Children tended to rate

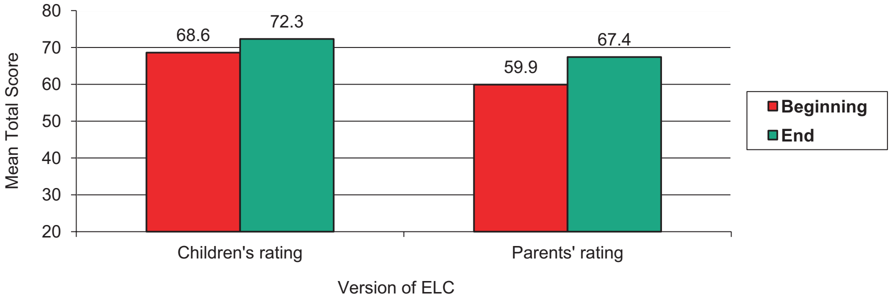

FIGURE 1 Average Total Scores on the Emotional Literacy Checklist, as rated by children ($n = 221$) and their parents ($n = 269$) at the beginning and end of therapy [Colour figure can be viewed at wileyonlinelibrary.com]

themselves higher than parents, though change in score was greater in parents. Positive changes in scores occurred across all the subscales scores also.

Table 2a details information on Parent-reported Emotional Literacy Checklist (ELC) scores, in terms of the proportions in different threshold bands. At the beginning of therapy, 199/269 (74.0%) of Parent-reported ELC scores were below or well below the average range and this changed to 126/269 (46.8%) by the end of therapy. Overall, scores across bands did not change for 120/269 (44.6%), improved for 133/269 (49.4%) and deteriorated for 16/269 (5.9%) indicating a statistically significant improvement in banding overall (Pearson Chi-square for trend = 85.1, $p < 0.001$).

Table 2b presents findings from the Child-reported Emotional Literacy Checklist (ELC) scores, in terms of the proportions in different threshold bands. At the beginning of therapy, 105/221 (47.5%) of scores were below or well below the average range and this changed to 74/221 (33.5%) by the end of therapy. Overall, scores across bands did not change or improved for 93/221 (42.1%), respectively,

TABLE 2 Proportions of participants in (a) Parent-reported Emotional Literacy Checklist score bands at the beginning and end of therapy ($n = 269$) and (b) Child-reported Emotional Literacy Checklist score bands at the beginning and end of therapy ($n = 221$)

		End of therapy Number (% of total)					
(a)		Well below average	Below average	Average	Above average	Well above average	Total
Beginning of therapy Number (% of total)	Well below average	72 (26.8)	32 (11.9)	43 (16.0)	5 (1.8)	1 (0.4)	153 (56.9)
	Below average	5 (1.8)	10 (3.7)	22 (8.2)	7 (2.6)	2 (0.7)	46 (17.1)
	Average	4 (1.5)	3 (1.1)	28 (10.4)	9 (3.3)	10 (3.7)	54 (20.1)
	Above average	0 (0)	0 (0)	3 (1.1)	5 (1.8)	2 (0.7)	10 (3.7)
	Well above average	0 (0)	0 (0)	1 (0.4)	0 (0)	5 (1.8)	6 (2.2)
	Total	81 (30.1)	45 (16.7)	97 (36.1)	26 (9.7)	20 (7.4)	269 (100.0)
Key	Worse	No change	Better				
		End of therapy (T2)					
(b)		Well below average	Below average	Average	Above average	Well above average	Total
Beginning of therapy (T1)	Well below average	26	16	19	1	2	64
	Below average	7	7	22	2	3	41
	Average	3	13	47	13	11	87
	Above average	1	1	7	7	4	20
	Well above average	0	0	2	1	6	9
	Total	37	37	97	24	26	221
Key	Worse	No change	Better				

(a) Cut-off scores: well below average = 60 or below; below average = 61–67; average = 68–80; above average = 81–86; well above average = 87 or above. (b) Cut-off for Pupil scores: well below average = 62 or below; below average = 63–68; average = 69–81; above average = 82–87; well above average = 88 or above. Cut-off for Student scores: well below average= 61 or below; below average = 62–66; average = 67–78; above average = 79–83; well above average = 84 or above and deteriorated for 35/221 (15.8%), indicating a statistically significant improvement in banding overall (Pearson Chi-square for trend = 12.43, $p < 0.01$).

Figure 2 details the average total and subscale scores for Outcome Rating Scale (ORS) self-report by young people. Again, total average scores improved, and this finding was statistically significant ($t = -8.1$, $p < 0.0001$). Improvements were apparent across all the subscales also. Table 3 presents the proportions of self-reported scores that crossed clinical and nonclinical thresholds. At the beginning of therapy 212/295 (71.9%) of scores were in the clinical range and this reduced to 148/295 (50.2) by the end of therapy. Overall, there was no change in clinical scores for 191/295 (64.7%), improvement for 84/295 (28.5%) and deterioration for 20/295 (6.8%). The percentage of participants' scores that exceeded the Reliable Change Index (RCI) of five points or more indicating significant improvement was 42.4% (126/297) and 11.1% (33/297) who showed significant deterioration.

Figure 3 details the average Total and Subscale scores for parent-reported Strengths and Difficulties Questionnaire (SDQ). Total average scores for the SDQ decreased, indicating improvement in functioning and this finding was statistically significant ($t = 7.0$, $p < 0.0001$), as were those for changes in the Subscale scores. Improvements were apparent across all the subscales also.

Table 4 presents the proportions of scores that crossed clinical thresholds. At the beginning of therapy 70/137 (51.1%) of scores were in the high/very high range and this reduced to 37/137 (27.0%) by the end of therapy. Overall, there was no band change in clinical scores for 57/137 (41.6%), improvement for 63/137 (46.0%) and deterioration for 17/137 (12.4%).

Table 5 provides a summary of the effectiveness of therapy. The Cohen's d statistic is used to measure effect size. Generally, an effect size of 0.2 is regarded as a small effect, 0.5 represents a medium effect and 0.8 a large effect size. The results indicate a medium effect size across most measures of well-being and functioning.

3.3 | Findings from experience of service questionnaires

Response rates to these questionnaires were lower (232/542; 42.8% for parents and 224/417; 53.7% for young people) and so must be treated with some caution. Figure 4 presents parents' feedback on the service received from CLEAR, using the Commission for Health Improvement Experience of Service Questionnaire (CHI-ESQ). Feedback was very positive across all aspects of the service. The convenient timing of appointments was rated the lowest. The mean Total score was 35, with 75.9% (154/203) of parents giving the service full marks (36/36).

Figure 5 presents young people's feedback on the service received, using the IHTPIW measure. Feedback was positive, and more varied. Three items were endorsed by at least 70% who completed the questionnaire:

- It's nice to speak to someone who isn't going to say anything to someone without permission.
- Just being able to chat has helped me to cope.
- I feel lots happier because I have someone I can talk to.

The item that was endorsed by the least number was:

- I understand more about my situation and I can connect with people around me.

The mean Total score for the IHTPIW was 35.7, with 13.5% (27/200) children and young people giving the service full marks (45/45).

Both experience of service questionnaires included open-ended questions which asked respondents to give feedback in the own words on what they liked about the service, what could be improved and any other matters. It is not possible to include analyses of these comments here. There were many specific positive and appreciative

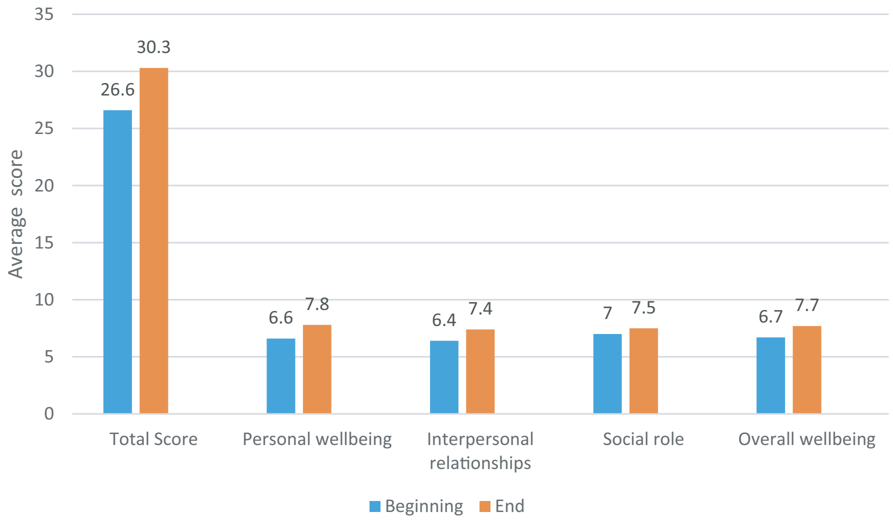

FIGURE 2 Average total and subscale scores for self-reported Outcome Rating Scale (ORS) and Child Outcome Rating Scale (CORS) at beginning and end of therapy ($n = 297$) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Proportions of participants in self-reported Outcome Rating Scale (ORS) and Child Outcome Rating Scale (CORS) score bands at the beginning and end of therapy (*n* = 295)

		End of therapy Number (% of total)		
		Clinical	**Non-clinical**	**Total**
Beginning of therapy Number (% of total)	Clinical	128 (43.4)	84 (28.5)	212 (71.9)
	Non-Clinical	20 (6.8)	63 (21.3)	83 (28.1)
	Total	148 (50.2)	147 (49.8)	295 (100.0)
Key	Worse	No change	Better	

Cut-off scores for self-reporting CORS = 32 and for ORS = 28.

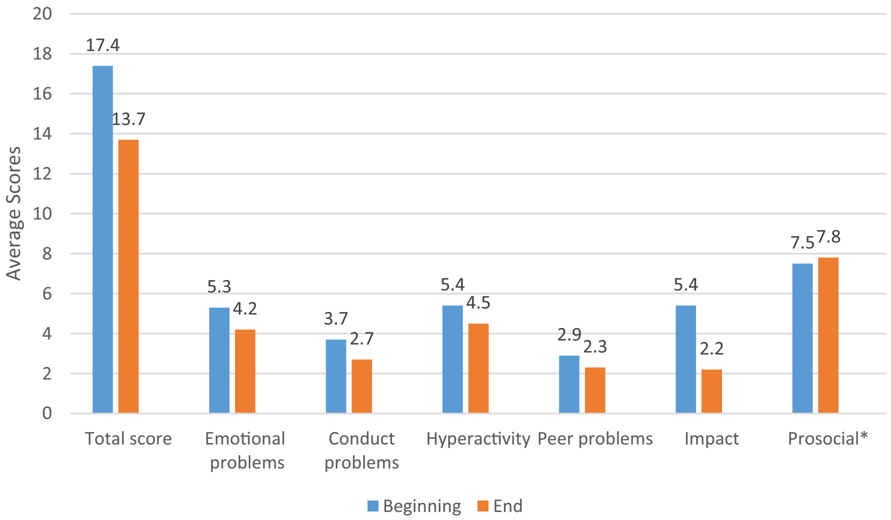

FIGURE 3 Average Total and Subscale Scores on parent-reported Strengths and Difficulties Questionnaire (SDQ), at the beginning and end of therapy (*n* = 137) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 4 Proportions of participants in Strengths and Difficulties Questionnaire Total score bands at beginning and end of therapy, as rated by Parents/Carers (*n* = 137)

		End of Therapy Number (% of total)				
		Close to average	**Slightly raised**	**High**	**Very high**	**Total**
Beginning of therapy Number (% of total)	Close to average	34 (24.8)	8 (5.8)	3 (3.6)	2 (1.5)	47 (34.3)
	Slightly raised	15 (10.9)	3 (2.2)	1 (0.7)	1 (0.7)	20 (14.6)
	High	8 (5.8)	6 (4.4)	2 (1.5)	2 (1.5)	18 (13.1)
	Very high	17 (12.4)	9 (6.6)	8 (5.8)	18 (13.1)	52 (38.0)
	Total	74 (54.0)	26 (19.0)	14 (10.2)	23 (16.8)	137 (100.0)
Key	Worse	No change	Better			

Cut-off scores: close to average = 0–13; slightly raised = 14–16; high = 17–19; and very high = 20–40.

comments. The most frequently reported way of improving the service was the availability of more therapy sessions.

4 | DISCUSSION

The most recent review of the evidence base in this field of practice identified no UK-based trials, few qualitative studies and little widespread service evaluation, with funding crises in this area of practice significantly undermining programme delivery (NIHR, 2016). This paper aims to address the very limited evidence base by presenting clinical effectiveness data on outcomes from individual therapy, and demonstrates what a small charity in the UK can achieve in this area of practice.

To summarise, children and young people who had experienced abusive relationships engaged well with individual therapy with over three quarters of all those referred completing at least three sessions of therapy. At least one complete outcome measure was

TABLE 5 Summary of therapeutic effectiveness

	Number	Mean beginning (*SD*)	Mean end (*SD*)	Cohen's *D* effect size
Emotional Literacy Checklist[a]				
Child and young person rated	221	68.5 (11.2)	72.3 (11.4)	0.34
Parent rated	269	59.9 (12.0)	67.4 (13.0)	0.60
Strengths and Difficulties Questionnaire				
Parent rated	134	17.4 (6.7)	13.7 (6.7)	0.55
Outcome Rating Scale/Child Outcome Rating Scale[a]				
Child and young person rated	297	26.6 (8.0)	30.3 (7.4)	0.48

[a] A higher score indicates better functioning.

available for nearly three quarters of the sample. Parents' and participants' experience of the service was generally positive, indicating that both found individual therapy acceptable.

Findings from the ROMs indicate that between one half and three quarters of children and young people referred to CLEAR had scores within the clinical or below average ranges. These participants presented with high needs. By the end of therapy, the proportion with high needs reduced to between one quarter and one half of scores. In terms of changes in thresholds score bands, there was improvement in scores for between 28.5 and 49.4% of self and parent-rated scores, a small proportion deteriorated (10% or less), but the biggest proportion of scores did not change. Overall, effect size, as measured by the ROMS, was small to medium. In the absence of a control group, it is not possible to confidently conclude that changes in well-being were a direct result of the intervention.

If these findings are compared with those achieved in the CYP-IAPT programme, it appears levels of engagement were higher (46% of closed cases had more than two recorded events); the availability of at least one complete outcome measure was similar (73% for CYP-IAPT); parent-reported ratings of the service indicated higher levels of satisfaction for all aspects of the service; proportionately, fewer children presented within the clinical range at the beginning of treatment (91% for CYP-IAPT); and trends for parent-reported and child-reported reliable deterioration, reliable improvement or no change were similar (Wolpert et al., 2016).

However, it is problematic to make direct comparisons when there are so many differences between the services, sample and

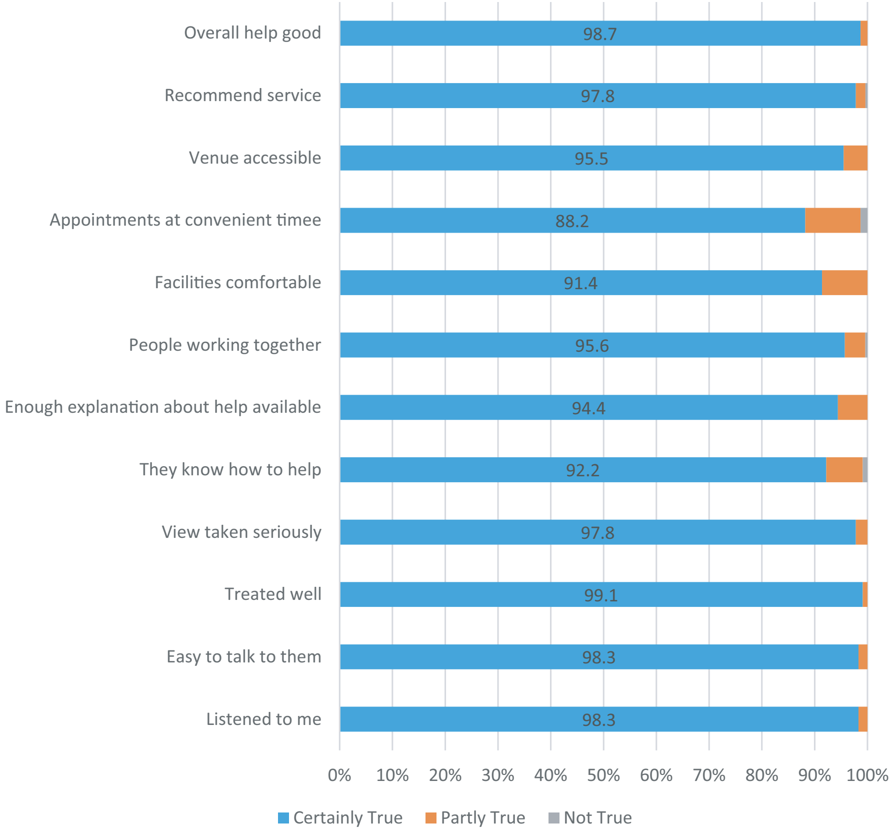

FIGURE 4 Parents' rating of the service as measured by the Commission for Health Improvement Experience of Service Questionnaire ($n = 232$) [Colour figure can be viewed at wileyonlinelibrary.com]

measures. Measuring outcomes and effectiveness is complex. The recent evaluation of the national CYP IAPT programme included 21 child-reported and 15 parent-reported outcome measures (Wolpert et al., 2016). The report reiterated that their evaluation was based on "flawed, uncertain, proximate and sparse" data; this is the nature of data in the real world (p7). In addition, how one synthesises, interprets and summaries a range of different measures in order to comment on "improvements" or "recovery" is also complicated.

Gathering evidence to demonstrate the effectiveness, or not, of therapy is fraught with difficulties. In this paper, we have tried to be careful about the language used to describe changes in outcome. Thus, statistical significance relates to scores on a ROM, rather than overall well-being. It is important not to mix up statistical significance with clinical significance. Thus, scores may change dramatically, but this does not necessarily mean that lives or well-being have also changed dramatically.

This paper has aimed to keep statistical analyses as descriptive as possible (for example, means and percentages in different thresholds) in order to avoid steps that require further defining, categorising and interpreting of raw data. We have reported on findings from each ROM separately, as there are real dilemmas in trying to combine outcomes across measures without making additional assumptions about the data and arbitrary decisions about what counts as real change. There is a danger that the more complicated the analyses, the further we move away from the voices of service users and their experience of therapy.

In the research world, clinical evaluation studies are sometimes regarded as inferior to gold standard randomised controlled trials (RCTs). However, evaluation studies may be better at capturing everyday practice and what can be realistically achieved in terms of measuring effectiveness in the "real world," rather than RCTs, which experimentally test the efficacy of an intervention in an "ideal world." We argue both types of design are necessary, but they address therapeutically and clinically distinct questions.

Our findings raise questions about the effectiveness of time limited individual therapy (average of 12 sessions per child) for those who have experienced abusive relationships. When there is a constant pressure to find funding for therapy, organisations (commissioners and service providers) often set limits based on economics (for instance, what can we afford) rather than what is in the best interests of the child. There is no robust evidence about how many sessions of therapy are effective for a child who has experienced significance violence and abuse. Indeed, it may not be possible to ever calculate such a number when so many factors need to be taken into account. Unfortunately, the more services are commissioned to provide certain sorts and amounts of therapy, the more difficult it is to evaluate the relative effectiveness of a diverse range of possible interventions.

Despite the drive towards evidence-based practice and practice-based evidence, the infrastructure and resources to do this well in the real world is extremely limited. We hope to have demonstrated what can be achieved within a small charity without access to dedicated research teams, academic departments or research grants.

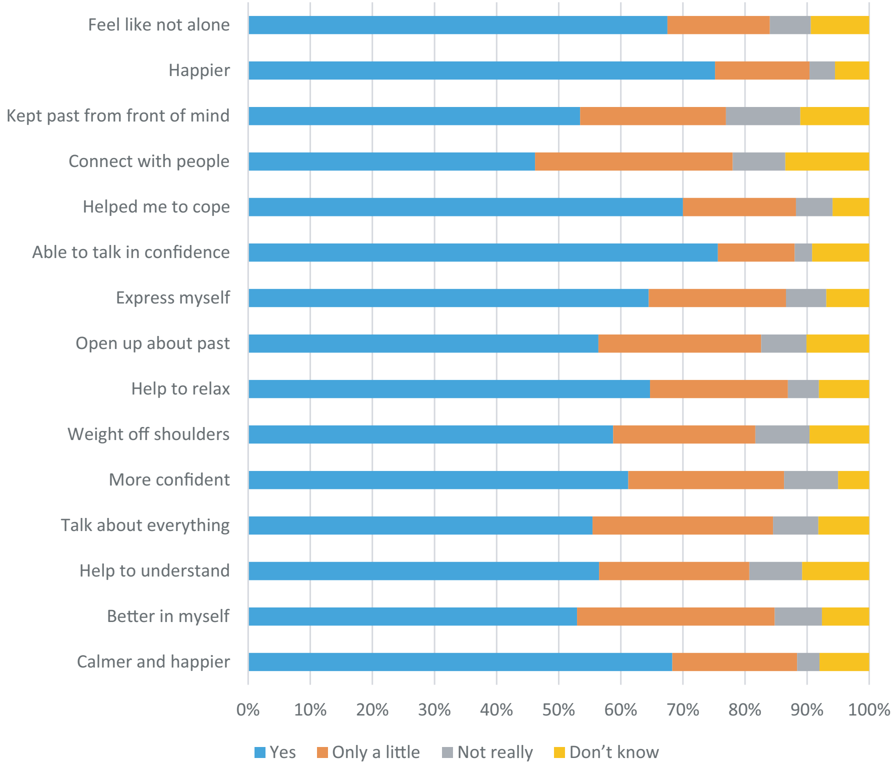

FIGURE 5 CYPs' rating of the service as measured by It's Hard to Put Into Words (n = 224) [Colour figure can be viewed at wileyonlinelibrary.com]

Setting up, reviewing and evaluating outcomes entailed contracting a consultant (the first author) for up to 1 day a week to carry out these functions. The charity also had to invest in information technology and support (databases and iPads) to collect ROMs. Completing ROMs also requires staff training and can be regarded as another factor that eats into precious time-limited therapeutic work rather than a necessary part of therapy. This is a significant investment that other charities may decide not to prioritise when finances are already stretched.

There are some significant constraints in the evaluation. Currently, therapists' views of the appropriateness and effectiveness of time-limited therapy and what additional support is needed has not been collected. A new questionnaire has been developed to capture this for each individual piece of work a therapist carries out. This may provide a better understanding of the strengths and limitations of current commissioning and provision arrangements.

In addition, the service was keen to use standardised measures, so that comparisons could be made with other studies and populations. However, in doing so, it has been subjected to the limitations of other studies and overlooked some basic questions. For example, participants and parents/carers were not asked directly to rate the effectiveness of therapy, the extent to which therapy changed their lives in ways that were important to them and what support they required to make long-lasting differences to their health and well-being.

Perhaps, involving young people and parents/carers more in the evaluation and research processes would have prevented these oversights. A recent report, "Making Noise: Children's voices for positive change after sexual abuse" (Warrington, Beckett, Ackerly, Walker & Allnock, 2017) illustrates how new research knowledge can be generated when the contributions of service users are prioritised. The participation of service users in evaluation and service-user-controlled research is limited (Faulkner, 2010) and this sort of exclusion may only serve to perpetuate the gaps in our knowledge about the effectiveness of therapy.

Although there were outcome data on nearly three quarters of the sample who completed therapy, better response rates may have been achieved if fewer or different ROMs were administered. We have prioritised the reporting of quantitative data over qualitative, perhaps because the former are taken more seriously than the latter in world of research and commissioning. Both are necessary to understand the processes that influence change. Embedding evaluation into a service and doing it well is problematic. If commissioners of services require good outcome data, additional investment in this aspect of practice is required. In our opinion, a focus on outcomes, without an understanding of the processes (therapeutic and evaluation), can create a climate of winners and losers rather than building resilient and creative services that work together to support the well-being of children and young people.

Finally, findings from this study may indicate that individual nondirective and time-limited therapy can improve well-being, as measured by the ROMs, but our research questions do not tell us whether the improvements achieved were sufficient, sustained and meaningful to the lives of service users and their families and how even better outcomes may be achieved, particularly for those whose scores did not change or got worse. Research does not have to deal with what happens after an intervention has been provided or withdrawn. In the real world, supporting the well-being of children and young people could be seen as an ongoing commitment by the whole community rather than a one-off event; a commitment that perhaps is not properly addressed in the current focus on evidence-based practice and commissioning arrangements, however well-intentioned.

ACKNOWLEDGEMENTS

The writing of this paper was dependent on the support of the CEO of CLEAR, Simon Carpenter, and all the therapists and service users who gave their time to administer and complete ROMs.

ORCID

Sarah Perry http://orcid.org/0000-0002-7482-0592

REFERENCES

Boeckel, M. G., Wagner, A., & Grassi-Oliveira, R. (2015). The effects of intimate partner violence on the maternal bond and PTSD symptoms of children. *Journal of Interpersonal Violence*, *32*, 1127–1142.

Bowlby, J. (1989). *The making and breaking of affectional bonds*. Abingdon, UK: Routledge.

British Columbia Centre of Excellence for Women's Health (BCCEWH) (2013). *Review of interventions to identify, prevent, reduce and respond to domestic violence*. Vancouver, BC: BCCEWH.

Brown, A., Ford, T., Deighton, J., & Wolpert, W. (2014). Satisfaction in child and adolescent mental health services. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*, 436–446.

Campbell, A., & Hemsley, S. (2009). Outcome rating scale and session rating scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*, *12*, 1–9. https://doi.org/10.1080/13284200802676391

Cater, A., & Øverlien, C. (2014). Children exposed to domestic violence: A discussion about research ethics and researchers' responsibilities. *Nordic Social Work Research*, *4*, 67–79. https://doi.org/10.1080/2156857X.2013.801878

Department for Education (2015). *Working together to safeguard children*. HM Government.

Department for Education (2016). *Characteristics of children in need: 2015 to 2106*. SFR52/2016. National Statistics.

Duncan, B. L., Sparks, J., Miller, S. D., Bohanske, R., & Claud, D. (2006). Giving youth a voice: A preliminary study of the reliability and validity of a brief outcome measure for children, adolescents and caretakers. *Journal of Brief Therapy*, *5*, 66–82.

Faulkner, A. (2010). *Changing our worlds. Examples of user-controlled research in action*. Eastleigh, UK: INVOLVE.

Faupel, A. (2003). *Emotional literacy. Assessment and intervention. Ages 7 to 11*. London, UK: NferNelson.

Felitti, V. J., Anda, R. F., Nordenberg Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE)

Study. *American Journal of Preventive Medicine*, *14*, 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Goodman, R. (2001). Psychometric properties of the strengths and difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Holt, S., Buckley, H., & Whelan, S. (2008). The impact of exposure to domestic violence on children and young people: A review of the literature. *Child Abuse & Neglect*, *32*(8), 797–810. https://doi.org/10.1016/j.chiabu.2008.02.004

Howarth, E., Moore, T. H. M., Welton, N. J., Lews, N., Stanley, N., MacMillan, H., … Feder, G. (2016). IMPRoving Oucomes for children exposed to domestic ViolencE (IMPROVE): An evidence synthesis. *Public Health Research*, *4*, 10.

Kitzmann, K. M., Gaylord, N. K., Holt, A. R., & Kenny, E. D. (2003). Child witnesses to domestic violence: A meta-analytic review. *Journal of Consulting & Clinical Psychology*, *71*(2), 339–352. https://doi.org/10.1037/0022-006X.71.2.339

Meltzer, H., Doos, L., Vostanis, P., Ford, T., & Goodman, R. (2009). The mental health of children who witness domestic violence. *Child Family Social Work*, *14*, 491–501. https://doi.org/10.1111/j.1365-2206.2009.00633.x

National Institute for Health and Care Excellence (2009). *Child maltreatment: When to suspect maltreatment in under 18s*. Retrieved from https://www.nice.org.uk/guidance/cg89

National Institute for Health and Care Excellence (2014). *Domestic violence and abuse: Multi-agency working*. London, UK: NICE.

National Institute for Health and Care Excellence (2017). *Child abuse and neglect*. https://www.nice.org.uk/guidance/ng76.

Perry, S., Carpenter, S., & CLEAR IDEAS (2016). Preliminary development and piloting of a user-generated routine outcome measure in a children and young people's counselling service. *Counselling and Psychotherapy Research*, *16*(3), 171–182.

Rizo, C. F., Macy, R. J., Ermentrout, D. M., & Johns, N. B. (2011). A review of family interventions for intimate partner violence with a child focus or child component. *Aggression Violent Behavior, 2011*(16), 144–166. https://doi.org/10.1016/j.avb.2011.02.004

Siegel, D. J. (1999). *The developing mind. How relationships and the brain interact to shape who we are*. New York, NY: Guildford Press.

Stanley, N. (2011). *Children experiencing domestic violence: A research review*. Dartington, UK: Research in Practice.

UNICEF (1989). *United Nations Convention on the Rights of the Child (UNCRC)*. London, UK: UNICEF.

Van der Kolk, B. (2015). *The body keeps the score. Mind, brain and body in the transformation of trauma*. London, UK: Penguin Random House.

Warrington, C., Beckett, H., Ackerly, E., Walker, M., & Allnock, D. (2017). *Making Noise: Children's voices for positive change after sexual abuse*. Bedfordshire, UK: University of Bedfordshire: NSPCC.

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and parent-reported outcomes and experience from child and young people's mental health services 2011–2015*. London, UK: CAMHS Press.

Women's Aid (2017). *Child first: A call to action one year on*. Bristol, UK: Women's Aid.

How to cite this article: Perry S, Frampton I. Measuring the effectiveness of individual therapy on the well-being of children and young people who have experienced abusive relationships, particularly domestic violence: A case study. *Couns Psychother Res*. 2018;18:356–368. https://doi.org/10.1002/capr.12184

BIOGRAPHIES

Sarah Perry is a Chartered and Clinical Psychologist. She has worked for statutory, voluntary and private sectors. She is an independent research consultant for CLEAR.

Ian Frampton is a Consultant in Paediatric Neuropsychology based in Cornwall UK, where his chief clinical interest concerns the development of children's brains. He is a Senior Lecturer at the Centre for Clinical Neuropsychology Research, University of Exeter and visiting Research Consultant at Oslo University Hospital, Norway.