



**EARLY
INTERVENTION
FOUNDATION**

MEASURING PARENTAL
CONFLICT AND ITS IMPACT
ON CHILD OUTCOMES

Measuring parental conflict and its impact on child outcomes

**Guidance on selecting and using valid,
reliable and practical measures to
evaluate interventions**

March 2020

Dr Inês Pote, Dr Virginia Ghiara, Dr Ekaterina Cooper, Laura Stock & Tom McBride

Acknowledgments

We are grateful to the Department for Work and Pensions for supporting this research. We would also like to thank those who contributed substantially to the contents of this report through their considerate comments and advice, including EIF colleagues Daniel Acquah, Kirsten Asmussen, Mark Ballinger, Elaine Drayton, Freyja Fischer, Ben Lewing and Thomas Masterman, as well as advisory group members, listed in appendix 4.

Disclaimer

This guide includes a selection of measures that could be used to assess if the outcomes of parents and children affected by parental conflict have improved after an intervention. It does not include an exhaustive list of all measures and is also not prescriptive, meaning that it does not recommend one particular measure.

The inclusion of a measurement tool in this guide does not necessarily constitute our endorsement of the tool. Instead, the guide provides an objective assessment of how valid, reliable and practical the included measures are. Due to time and resource constraints, the methodological approach used in this review did not involve an exhaustive search of all the literature underpinning a measure. There is therefore a risk that we have missed relevant evidence when applying our measurement assessment criteria, meaning that our assessment of each measure may be incomplete. We also did not involve the measurement developers in this work, although we did contact some developers where we felt it was necessary.

Finally, we have only assessed one version of each selected measure, with the exception of the Dyadic Adjustment Scale, where we have assessed the DAS-7 and DAS-32. In the summary and full measurement reports we have listed all the available versions identified from our review of the literature. It should not be assumed that other versions of the same measure would yield an equal rating to that presented in this guide.

About EIF

The Early Intervention Foundation (EIF) is an independent charity established in 2013 to champion and support the use of effective early intervention to improve the lives of children and young people at risk of experiencing poor outcomes.

Effective early intervention works to prevent problems occurring, or to tackle them head-on when they do, before problems get worse. It also helps to foster a whole set of personal strengths and skills that prepare a child for adult life.

EIF is a research charity, focused on promoting and enabling an evidence-based approach to early intervention. Our work focuses on the developmental issues that can arise during a child's life, from birth to the age of 18, including their physical, cognitive, behavioural and social and emotional development. As a result, our work covers a wide range of policy and service areas, including health, education, families and policing.

EIF IS PROUD TO BE A MEMBER OF
THE WHAT WORKS NETWORK



Early Intervention Foundation
10 Salamanca Place
London SE1 7HB

W: www.EIF.org.uk
E: info@eif.org.uk
T: [@TheEIFoundation](https://twitter.com/TheEIFoundation)
P: +44 (0)20 3542 2481

This paper was first published in March 2020. © 2020

The aim of this report is to support policymakers, practitioners and commissioners to make informed choices. We have reviewed data from authoritative sources but this analysis must be seen as a supplement to, rather than a substitute for, professional judgment. The What Works Network is not responsible for, and cannot guarantee the accuracy of, any analysis produced or cited herein.

EIF is a registered charity (1152605) and a company limited by guarantee (8066785).

Download

This document is available to download as a free PDF at:
<https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

Permission to share

This document is published under a creative commons licence: Attribution-NonCommercial-NoDerivs 2.0 UK
<http://creativecommons.org/licenses/by-nc-nd/2.0/uk/>



For commercial use, please contact info@eif.org.uk

Summary

Conflict between parents – whether together or separated – is a normal part of relationships. However, as set out in our 2016 review in collaboration with Professor Gordon Harold, there is a strong body of evidence suggesting that **when parental conflict is frequent, intense and poorly resolved, it can have a significant negative impact on children's outcomes.**¹

Supporting healthy relationships between parents is therefore an increasingly important focus for early intervention and local family services. However, it is not always easy to select child and family outcome measures that will tell you if an intervention is making a difference to the lives of families.

This practical guide has been developed to improve understanding of **how to measure if the outcomes of the parents and children affected by parental conflict have improved after an intervention.** The guide identifies measures that can be used to assess progress for individual families, and explores how valid, reliable and practical these tools are. While the guide is framed with reducing parental conflict in mind, it could also be of value to those involved with children and family services more broadly.

There are many measures that could be used. **This guide includes a selection of some of the most relevant and frequently used measures rather than an exhaustive list.** It is also not a prescriptive guide, meaning that it does not recommend one particular measure. The outcomes we focus on are informed by the Family Stress Model, which was developed by Professor Gordon Harold, building on the work of Conger and colleagues.² This model shows that the quality of the interparental relationship (including parental conflict) is a central mechanism through which social and contextual family pressure (such as economic pressure) can impact on both parenting and children's outcomes.

Child outcomes include:

- **Internalising behaviours**, characterised by symptoms of withdrawal, inhibition, fearfulness, sadness, shyness, low self-esteem, anxiety, depression and suicidality in its most extreme.
- **Externalising behaviours**, characterised by a broad set of behavioural difficulties, including aggression, hostility, non-compliant and disruptive behaviours, verbal and physical violence, anti-social behaviour, conduct disorder, delinquency and even vandalism.

Interparental (or couple) relationship outcomes include:

- **Relationship quality**, including factors around couple satisfaction, commitment, communication, respect, consensus and shared interests.
- **Parental conflict**, specifically the frequency and intensity of conflict, conflict resolution skills, and children's perception of and adjustment to the conflict between parents.
- **Co-parenting practices**, including parents' ability to cooperate and communicate when performing parenting duties.

This guide includes 18 measures, seven assessing child outcomes and 11 assessing interparental (or couple) relationship outcomes. Each measure has been assessed against EIF criteria for how valid, reliable and practical they are.

¹ Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance inter-parental relationships and improve outcomes for children*. Early Intervention Foundation: London. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>

² Conger, R.D., Ge, X., Elder, G.H., Lorenz, F.O., & Simons, R.L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, 65(2), 541–561.

Summary of assessments:

Child outcome measures

| | Child measures Outcomes assessed CLICK ON A MEASURE FOR MORE DETAIL | Respondent Who should complete this version of the measure? | Target population: Who is the measure developed for? | Psychometric features: How valid and reliable is the measure? | Implementation features: How practical is the measure? |
|--|--|--|---|--|---|
| Internalising & externalising behaviours | Child Behaviour Checklist (CBCL/6–18) This measure assesses the child's emotional and behavioural problems, including (but not limited to) anxiety, depression, hyperactivity and aggressive behaviour. | Parents of children aged 6–18 years | Children aged 6–18 years | 4/4 | 1/4 |
| | Strengths and Difficulties Questionnaire (SDQ) This measure assesses the child's emotional and behavioural problems, including (but not limited to) prosocial behaviour, hyperactivity, conduct and peer problems. | Parents of children aged 4–17 years | Children aged 4–17 years | 4/4 | 4/4 |
| Internalising behaviours | Mood and Feelings Questionnaire (MFQ) This measure assesses the child's feelings, thoughts and behaviours associated with depression. | Children aged 8–18 years | Children aged 8–18 years | 4/4 | 4/4 |
| | Revised Child Anxiety and Depression Scale (RCADS) This measure assesses the child's symptoms of depressive and anxiety disorders. | Children aged 8–18 years | Children aged 8–18 years | 3/4 | 3/4 |
| | Screen for Child Anxiety Related Emotional Disorders (SCARED) This measure assesses the child's symptoms of different anxiety disorders. | Children aged 8–18 years | Children aged 8–18 years | 3/4 | 3/4 |
| Externalising behaviours | Eyberg Child Behavior Inventory (ECBI) This measure assesses the frequency and severity of disruptive behaviours manifested by the child. | Parents of children aged 2–16 years | Children aged 2–16 years | 3/4 | 2/4 |
| | Parent Daily Report (PDR) This measure assesses negative behaviours displayed by the child at home. | Parents of children aged 4–10 years | Children aged 4–10 years | 3/4 | 3/4 |

Note: Measures have been rated according to the assessment criteria presented in appendix C, table C.2. The colour code has been based on the following ratings: 4/4 = green; 3/4 = grey; and 1/4 or 2/4 = red. Measures for which we obtained insufficient evidence to establish validity (for example PDR) have also been labelled red, irrespective of their total scoring, as validity is considered to be one of the most important psychometric features of a measure.

Summary of assessments:

Interparental (or couple) relationship outcome measures

| | Interparental (couple) relationship measures Outcomes assessed CLICK ON A MEASURE FOR MORE DETAIL | Respondent: Who should complete this version of the measure? | Target population: Who is the measure developed for? | Psychometric features: How valid and reliable is the measure? | Implementation features: How practical is the measure? |
|----------------------|--|--|--|---|--|
| Relationship quality | Couples Satisfaction Index (CSI-16) This measure assesses how satisfied a person is in their relationship and how they feel about it. | Adults in a relationship | Intact couples | 2/4 | 4/4 |
| | Dyadic Adjustment Scale (DAS-32) This measure assesses how satisfied a person is in their relationship, the feelings associated with the relationship, and the issues causing disagreements between partners. | Adults in a relationship | Intact couples | 3/4 | 4/4 |
| | Dyadic Adjustment Scale (DAS-7) This measure assesses how satisfied a person is in their relationship and the extent of agreement or disagreement among the couple over important aspects of life. | Adults in a relationship | Intact couples | 3/4 | 4/4 |
| | Golombok Rust Inventory of Marital State (GRIMS) This measure assesses how a person feels about their own and their partner's behaviour within their relationship as well as their attitudes and feelings about the relationship. | Adults in a relationship | Intact couples | 2/4 | 4/4 |
| | Marital Adjustment Test (MAT) This measure assesses how satisfied a person is in their marital relationship and the extent of agreement or disagreement among the couple over important aspects of life. | Adults who are married | Married couples | 3/4 | 4/4 |
| | Relationship Quality Index (RQI) This measure assesses how satisfied a person is in their relationship and the extent of agreement or disagreement among the couple over important aspects of life. | Adults in a relationship | Intact couples | 3/4 | 3/4 |
| Parental conflict | Children's Perception of Interparental Conflict Scale (CPIC) This measure assesses parental conflict from the child's point of view, particularly in terms of the frequency, intensity, resolution and perceived threat of the conflict. It also focuses on how the child responds to the conflict, including questions around self-blame and coping mechanisms. | Children aged 9–17 years with intact or separated parents | Intact and separated couples with children | 2/4 | 3/4 |
| | O'Leary Porter Scale (OPS) This measure assesses the frequency of couple hostility observed by the child, including quarrels, sarcasm and physical abuse. | Parents in a relationship | Intact couples with children | 2/4 | 4/4 |
| Co-parenting | Parenting Alliance Measure (PAM) This measure assesses how cooperative, communicative and mutually respectful parents are when caring for their children. | Parents who are in a relationship or are separated | Intact and separated couples with children | 2/4 | 3/4 |
| | Parent Problem Checklist (PPC) This measure assesses the extent of agreement or disagreement between the parents over child-rearing issues. | Parents who are in a relationship or are separated | Intact and separated couples with children | 2/4 | 4/4 |
| | Quality of Co-parental Communication Scale (QCCS) This measure assesses the extent of mutual support and hostility over child-rearing issues with the former spouse. | Parents who are separated | Separated couples with children | 2/4 | 3/4 |

Note: Measures have been rated according to the assessment criteria presented in appendix C, table C.2. The colour code has been based on the following ratings: 4/4 = green; 3/4 = grey; and 1/4 or 2/4 = red. Measures for which we obtained insufficient evidence to establish validity (for example MAT) have also been labelled red, irrespective of their total scoring, as validity is considered to be one of the most important psychometric features of a measure. Please note that many of these measures have items assessing more than one of the three outcomes listed on the left-hand side column. Measures have therefore been grouped according to the outcome most prevalent within the measure.

More information on each measure can be found within the individual **summary reports**, included in appendix A. The **full measurement reports**, including details of the evidence used to assess each measure against our predefined psychometric and implementation criteria, are supplementary to this report and can be downloaded from the EIF website.³ We strongly encourage you to engage with this more detailed evidence so that you understand the strengths and limitations of the measures you are considering, and can make more informed and evidence-based decisions.

We conclude the guide by providing readers with some **key recommendations on how best to select and use outcome measures**, including practical tips to consider. In terms of selecting appropriate measures, for example, we encourage users to:

- select validated measures that are consistent with the intervention's theory of change and capable of assessing at least one of the intended outcomes
- consider measures that are appropriate for the target population under consideration, paying special attention to participant age, level of need, demographic characteristics and preferred language
- select an adequate number and appropriate combination of measures that will not overburden participants and will take into consideration the time and resources available for the evaluation, as well as the cost, licensing, ease of scoring and training requirements of each measure
- consider including measures completed by different respondents and/or collect administrative or observational data, if interested in conducting more robust and ambitious evaluations
- conduct further due diligence by consulting with subject-matter experts, to decide whether the selected measures are the most suitable ones.

We also provide some guidance on how to use measures appropriately, suggesting readers:

- ensure measures are completed at least twice, once before participants receive the intervention and then again after the intervention has been delivered
- use measures in their entirety, without changing, adding, deleting or altering the order of any items
- administer measures in a standardised way, to ensure data is collected validly and reliably
- consider logistical aspects of research, including ethical approval, informed consent and data collection, as set out by the General Data Protection Regulation (GDPR)
- introduce measures in a sensitive manner and respond appropriately if the data collected highlights a particular safeguarding risk
- use the information collected from participants to assess whether the intervention or service has worked
- acknowledge the authors and developers of the measures used
- be aware of the strengths and limitations of the measures used.

Ultimately, the guide is intended to support improvements to the quality of evaluation, contributing to more high-quality evidence on the effectiveness of early intervention programmes in the UK.

³ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

Contents

| | |
|---|-----|
| Summary | 3 |
| 1. Introduction | 8 |
| 1.1 Why is this guide needed? | 8 |
| 1.2 Who is this guide for? | 9 |
| 1.3 The context behind this guide | 9 |
| 2. A brief summary of the evidence on parental conflict | 12 |
| 2.1 Why reducing parental conflict matters for children | 12 |
| 2.2 Why we need to improve how we measure child outcomes | 14 |
| 2.3 Outcomes we are focusing on in this guide | 15 |
| 3. Why measure? | 17 |
| 3.1 What are validated measures and why is using them important? | 18 |
| 3.2 Additional features of desirable measures | 21 |
| 4. Measures | 22 |
| 4.1 Selecting relevant measures | 22 |
| 4.2 Assessing selected measures against our psychometric and implementation criteria | 25 |
| 4.3 A snapshot of the validity, reliability and practicality of our selected measures | 25 |
| 5. Guidance on selecting and using outcome measures | 34 |
| 5.1 Selecting appropriate measures | 34 |
| 5.2 Using measures appropriately | 41 |
| 5.4 Helpful resources | 45 |
| Appendix A: Summary reports of outcome measures | 48 |
| Child outcome measures | |
| Child Behaviour Checklist (CBCL/6–18) | 49 |
| Strengths and Difficulties Questionnaire (SDQ) | 52 |
| Mood and Feelings Questionnaire (MFQ) | 55 |
| Revised Child Anxiety and Depression Scale (RCADS) | 58 |
| Screen for Child Anxiety Related Emotional Disorders (SCARED) | 61 |
| Eyberg Child Behavior Inventory (ECBI) | 64 |
| Parent Daily Report (PDR) | 67 |
| Interparental outcome measures | |
| Couples Satisfaction Index (CSI-16) | 70 |
| Dyadic Adjustment Scale (DAS-32) | 73 |
| Dyadic Adjustment Scale (DAS-7) | 76 |
| Golombok Rust Inventory of Marital State (GRIMS) | 79 |
| Marital Adjustment Test (MAT) | 82 |
| Relationship Quality Index (RQI) | 85 |
| Children's Perception of Interparental Conflict Scale (CPIC) | 88 |
| O'Leary Porter Scale (OPS) | 91 |
| Parenting Alliance Measure (PAM) | 94 |
| Parent Problem Checklist (PPC) | 97 |
| Quality of Co-parental Communication Scale (QCCS) | 100 |
| Appendix B: Guiding you to identify measures used in impact evaluations | 103 |
| Appendix C: Methodology | 109 |
| Appendix D: Advisory members | 118 |
| References | 119 |

1. Introduction

1.1 Why is this guide needed?

Conflict between parents – whether together or separated – can have a significant negative impact on effective parenting and children’s outcomes.^{1,2} Supporting healthy relationships between parents is therefore an increasingly important focus for early intervention and local family services. **You want to know that what you are delivering is improving the lives of the families you are seeking to help.** However, measuring whether the impact of the interventions you deliver are making a difference, is difficult.

We know that measuring child and family outcomes is one of the most challenging issues that local areas grapple with when delivering parental conflict programmes. In fact, one of the most common questions we get asked for help with is how to measure outcomes effectively. The problem is that there are a wide variety of tools and outcome measures available, each capturing an array of different things, some more suitable for your intervention than others.

The key thing to keep in mind here is that, **to understand the impact an intervention is having, you need to measure change in your outcomes of interest using valid and reliable instruments.** In the case of an intervention seeking to reduce parental conflict, for example, the outcomes of interest may be the frequency and intensity of the conflict between parents as well as the child’s behaviour. However, if you were to use invalid and unreliable measures of these outcomes, it would not be possible to determine whether any apparent improvement in outcomes was due to a true effect of the intervention under evaluation, or a product of measurement error and bias. Without valid and reliable measures, it is therefore not possible to have confidence in the findings and conclusions of studies.

The purpose of this guide is to help you understand **how to measure if the outcomes of the parents and children affected by parental conflict have improved after an intervention.** It identifies examples of measures that could be used to assess progress for individual families and describes how valid, reliable and practical these tools are. Importantly though, **this guide does not include an exhaustive list of all the measures** you could consider. It simply includes a selection of some of the most relevant and frequently used measures in this space. **It is also not a prescriptive guide,** meaning that it does not recommend one particular measure. Instead, it explains why measuring outcomes is so important and what good measurement actually looks like, providing you with guidance on how to select and use the most appropriate measures for your situation.

This guide builds on and expands our previous work in this area, including the 2016 review on *What works to enhance inter-parental relationships and improve outcomes for children*,³ led by Professor Gordon Harold from the University of Sussex. In order to fully understand how parental conflict can impact on children and how programmes that aim to enhance the interparental relationship can support improved outcomes, we encourage all our readers to engage fully with this body of literature before embarking on the delivery and evaluation of programmes in this area.

¹ Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance inter-parental relationships and improve outcomes for children*. Early Intervention Foundation: London. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>

² Harold, G., and Sellers, R. (2018). Annual Research Review: Interparental conflict and youth psychopathology: an evidence review and practice focused update. *Journal of Child Psychology and Psychiatry* 59:4, 374–402.

³ Harold et al. (2016). Op. cit.

1.2 Who is this guide for?

This guide has mainly been developed for those involved in commissioning, planning and delivering evaluations. It contains some technical content and assumes an understanding of how parental conflict impacts children, the outcomes that programmes targeting the interparental relationship seek to achieve, and approaches to evaluation.

While the guide is framed with reducing parental conflict in mind, we think it could also be of value to those involved with children and family services more broadly, including parenting interventions as well as child and adolescent mental health services.

Ultimately, our aim is that this guide will support improvements to the quality of evaluation, contributing to more high-quality evidence on the effectiveness of early intervention programmes in the UK.

Interested in learning more about evaluation?

Conducting a high-quality evaluation which seeks to determine intervention effectiveness is a challenging thing to do – and measurement is only a small part of that challenge. Therefore, if you are interested in learning about how to evaluate individual programmes or services more generally, beyond just measurement, we suggest you have a look at three of the Early Intervention Foundation's (EIF) most recent evaluation reports.

- **10 steps for evaluation success** helps readers evaluate individual programmes by guiding them through all the stages of evaluation maturity, from theory of change to impact assessments, and the quality assurance systems that are essential if interventions are going to remain effective when offered at scale.⁴
- **Evaluating early intervention programmes: Six common pitfalls, and how to avoid them** provides guidance on addressing some of the most common issues we see in our assessments of programme evaluations, including explanations of how these problems undermine confidence in a study's findings, and how they can be avoided or rectified.⁵
- **Evaluating early help** provides advice on how to apply the principles of good impact evaluation to complex local systems, such as an early help offer.⁶

1.3 The context behind this guide

Policy and practice context

Improving children's outcomes by reducing parental conflict has become an important focus in policy and practice over recent years. In 2017 the Department for Work and Pensions (DWP) launched the *Improving Lives*⁷ strategy to tackle the multiple disadvantages experienced by workless families, including parental conflict. This followed our 2016 review, led by Professor Gordon Harold, which presented robust research evidence of how conflict between parents, specifically how they communicate and relate to each other, can have a significant and detrimental impact on effective parenting, children's long-term mental health and their future life chances.⁸

⁴ See: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

⁵ See: <https://www.eif.org.uk/resource/evaluating-early-intervention-programmes-six-common-pitfalls-and-how-to-avoid-them>

⁶ See: <https://www.eif.org.uk/resource/evaluating-early-help-a-guide-to-evaluation-of-complex-local-early-help-systems>

⁷ Department for Work and Pensions. (2017a). *Improving Lives: Helping Workless Families*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/621364/improving-lives-helping-workless-families-web-version.pdf

⁸ Harold et al. (2016). Op. cit.

Over the past few years, national policy has evolved from an emphasis on family stability and family structure,⁹ to tackling how children are affected by their parents' relationship, regardless of whether they are together or separated.¹⁰ There has also been increased attention on improving child mental health, as in the 2017 green paper on *Transforming children and young people's mental health provision*. This paper highlighted how children exposed to persistent and unresolved parental conflict are at greater risk of emotional and behavioural problems in both the short and long term.¹¹ Disadvantaged families and those in poverty are also more at risk,¹² with DWP analysis estimating that relationship distress is three times more prevalent in workless families.¹³ Given this context, DWP launched the national Reducing Parental Conflict Programme in 2017, investing £39 million until 2021, to embed efforts to address parental conflict at a local level, and improve both the supply and demand for evidence-based interventions. Similarly, the current phase of the Troubled Families Programme includes a greater emphasis on addressing parental conflict as a key issue for families with complex needs.

Our work at EIF

Building on our 2016 report, we have undertaken a range of research and practical projects to support local areas in reducing parental conflict,¹⁴ including our 2017 evidence review on how poverty and economic hardship can impact on the quality of the interparental relationship and on children's outcomes.¹⁵

More recently, we have been working with DWP on the national Reducing Parental Conflict Programme. This has involved developing practical tools to support the design and delivery of local parental conflict services, such as a commissioner's guide¹⁶ and a planning tool.¹⁷ Although both these resources refer to the importance of measuring outcomes for families taking part in programmes targeting the interparental relationship, they do not help local areas decide what measures to select. This is why we have developed this guide – to provide an objective assessment of the psychometric and implementation features of some of the measures that could be used to assess the impact of these interventions.

As an independent UK What Works centre, EIF is a leading authority on evaluation and how to know whether interventions 'work'. Our flagship product, the EIF Guidebook,¹⁸ provides information on the strength of evidence underpinning over 100 early intervention programmes that have been shown to improve outcomes for children and young people, including a number of programmes targeting interparental relationships.¹⁹ We also have

⁹ Department for Children Schools and Families. (2010). *Support for all: the Families and Relationships Green Paper*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/329448/Support_for_All.pdf

¹⁰ Department for Work and Pensions. (2017a). Op. cit.

¹¹ Department of Health and Social Care and Department for Education. (2017). *Transforming Children and Young People's Mental Health Provision: A Green Paper*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/664855/Transforming_children_and_young_people_s_mental_health_provision.pdf

¹² Acquah, D., Sellers, R., Stock, L., & Harold, G. (2017). *Inter-parental conflict and outcomes for children in the contexts of poverty and economic pressure*. London: Early Intervention Foundation. <https://www.eif.org.uk/report/interparental-conflict-and-outcomes-for-children-in-the-contexts-of-poverty-and-economic-pressure>

¹³ Department for Work and Pensions. (2017b). *Improving Lives: Helping Workless Families Analysis and Research Pack*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696368/improving-lives-helping-workless-families-web-version.pdf

¹⁴ See: <https://reducingparentalconflict.eif.org.uk/>

¹⁵ Acquah et al. (2017). Op. cit.

¹⁶ See: <https://www.eif.org.uk/resource/commissioner-guide-reducing-parental-conflict>

¹⁷ See: <https://www.eif.org.uk/resource/reducing-parental-conflict-planning-tool>

¹⁸ See: <https://guidebook.eif.org.uk/>

¹⁹ See: <https://guidebook.eif.org.uk/search?sets%5B%5D=%25%22improving-interparental-relationships%22%25>

relevant expertise on how to collect robust evidence that an intervention or service has made a meaningful and measurable difference to the lives of children, and have produced several hands-on guides to evaluation.²⁰

A note on language

Throughout this guide, we refer to parents and couples interchangeably, mentioning both the interparental and couple relationship. When we use either of these terms, we are being as inclusive as possible, referring to any intact or separated couple with or without children.

²⁰ For more information on evaluation, see the box in section 1.2.

2. A brief summary of the evidence on parental conflict

2.1 Why reducing parental conflict matters for children

Conflict between parents is a normal part of relationships. However, as set out in our 2016 review in collaboration with Professor Gordon Harold, there is a strong body of evidence suggesting that when parental conflict is frequent, intense and poorly resolved, it puts children's mental health and long-term outcomes at risk, irrespective of whether parents are together or separated, in step, foster or blended families.²¹ These destructive conflict behaviours include aggression, non-verbal conflict or the 'silent treatment', lack of respect and emotional control, lack of resolution and, in their most extreme form, domestic abuse.

Disclaimer: a note on domestic abuse

Although destructive parental conflict can include domestic abuse, the focus of this report and prior EIF reports on the topic of parental conflict has been on non-abusive conflict between parents that is frequent, intense and poorly resolved. We have therefore not included measures of domestic abuse within this guide, despite these existing.²² We do, however, appreciate that the quality of interparental relationships is dynamic in nature, and so it would be possible for some couples enrolled in a parental conflict intervention to start off in a non-abusive relationship which escalates into abuse, despite involvement in the programme. Limiting outcome measures to only assess parental conflict without assessing the potential presence of domestic abuse, risks missing an important part of the overall picture.

It is the responsibility of the intervention provider to carry out a preliminary domestic abuse risk assessment and to determine the couple's suitability to take part in the intervention.

The risk of conflict is also higher at crucial transition points in family life, such as becoming pregnant, having a baby, children starting or changing school, and separation or divorce. Despite this, it is important to note that children of all ages can be affected by destructive parental conflict, with effects evidenced across infancy, childhood, adolescence and adulthood.²³

²¹ Harold et al. (2016). Op. cit.

²² See, for example, the Revised Conflict Tactics Scale (CTS-2; Straus, M.A., Hamby, S.L., Boney-McCoy, S., and Sugarman, D.B. (1996). The Revised Conflict Tactics Scale (CTS2): development and preliminary psychometric data. *Journal of Family Issues* 17, 283-316. Available at: <https://journals.sagepub.com/doi/10.1177/019251396017003001>) and the Composite Abuse Scale (CAS; Hegarty, K., Sheehan, M., and Schonfeld, C. (1999). A multidimensional definition of partner abuse: development and preliminary validation of the Composite Abuse Scale. *Journal of Family Violence* 14, 399-415. Available at: <https://link.springer.com/article/10.1023/A:1022834215681>).

²³ Harold et al. (2016). Op. cit.

According to longitudinal evidence reviewed in the 2016 report, parental conflict can lead to a wide variety of difficulties, including:

- **emotional and behavioural difficulties**, involving internalising and externalising problems
 - where **internalising problems** are characterised by symptoms of withdrawal, inhibition, fearfulness, sadness, shyness, low self-esteem, anxiety, depression and suicidality in its most extreme
 - and **externalising problems** are characterised by a broad set of behavioural difficulties, including aggression, hostility, non-compliant and disruptive behaviours, verbal and physical violence, anti-social behaviour, conduct disorder, delinquency and even vandalism
- **academic problems**, including low academic attainment and later employment difficulties
- **physical health problems**, such as higher rates of illness and fatigue as well as impacts on risky health behaviours such as smoking, substance misuse and early sexual activity
- **social and interpersonal relationship problems**, including more hostile relationships with siblings, conflicts with peers and difficulties with future romantic relationships.

Importantly, there is also evidence to suggest that these outcomes can converge and accumulate over an individual's childhood and adolescence to cause significant reductions in their overall life chances, as well as risk relationship problems and behavioural patterns being repeated across generations. A child exposed to frequent, intense and poorly resolved parental conflict, for example, is at increased risk of internalising and externalising problems, which in turn may lead to more negative academic outcomes, peer conflicts, substance misuse, low employability, reduced couple and parenting proficiency, and future disrupted family and child outcomes.²⁴

Another key research finding highlighted in both the 2016 and 2017 evidence reviews is that the quality of the interparental relationship is a primary influence or central mechanism through which family stress can impact on both parenting and children's outcomes. According to the internationally replicated Family Stress Model (figure 2.1), social and contextual pressure (such as economic hardship) can impact on parental mental health, which in turn can affect the quality of the interparental relationship and increase parental conflict. Conflict between parents can then directly lead to poor outcomes for children, or it can result in the use of harsh and inappropriate parenting practices (including poor co-parenting in separated couples), which can affect the quality of the parent-child relationship and negatively impact on child outcomes.²⁵

Research reviewed as part of the 2016 report also highlighted two primary processes through which parental conflict has been shown to affect children. According to one of these processes, it is hypothesised that the effects of conflict between parents are deemed to occur indirectly through a 'spillover' of emotion from the couple relationship to the parent-child relationship, whereby parents in distressed relationships tend to be more hostile and aggressive towards their children as well as less sensitive and emotionally responsive to their children's needs.²⁶ Research demonstrating that overt parental conflict to which children are exposed has a greater impact on child distress than covert conflict to which children are not exposed,²⁷ has led some academics to consider another primary process through which parental conflict affects children. According to this process, which emerged from three

²⁴ Ibid

²⁵ Conger, R.D., Ge, X., Elder, G.H., Lorenz, F.O., & Simons, R.L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, 65(2), 541–561.

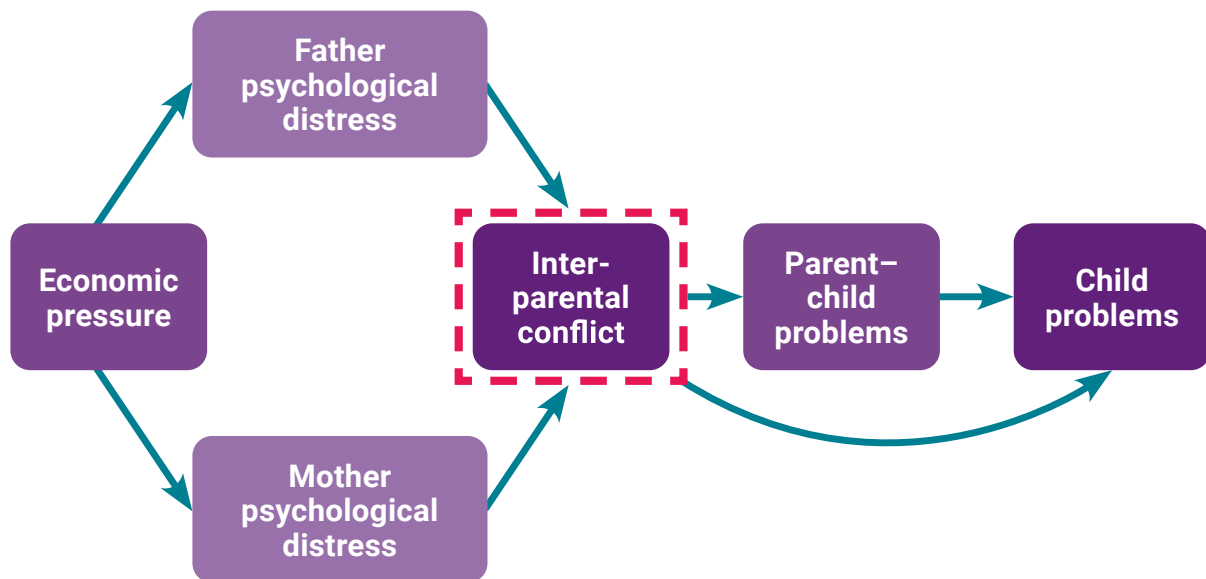
²⁶ Erel, O., & Burman B. (1995). Interrelatedness of marital relations and parent-child relations: A meta-analytic review. *Psychological Bulletin*, 118(1), 108–132.

²⁷ Cummings, E., & Davies, P. (2002). Effects of marital conflict on children: Recent advances and emerging themes in process-oriented research. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(1), 31–63.

theoretical perspectives, it is the negative emotions, cognitions and representations of family relationships engendered in children who live in acrimonious households, which explain the effect of parental conflict on children's poor psychological outcomes.²⁸

FIGURE 2.1

Family Stress Model



Source: Harold et al. (2016)

Considering this, it is important that parenting interventions targeting families in which high levels of parental conflict are likely, include a focus on the quality of the interparental relationship in order to maximise their chances of success.²⁹

2.2 Why we need to improve how we measure child outcomes

Despite the strong scientific evidence of how parental conflict impacts on children, the UK evidence of 'what works' to effectively address parental conflict and improve outcomes for children is still at an early stage.³⁰ **This means we don't yet know what impact we are having on families and on children** in particular.

Most programmes and services targeting the quality of interparental relationships **do not measure if and how they improve child outcomes** – and if they do collect evidence on impact, the majority only do this for parents and couples. Of the 15 UK interventions that went through the EIF programme assessment process as part of the 2016 What Works review, only one was found to have preliminary evidence of improving child outcomes.³¹ Considering that many of these programmes were not designed with child outcomes as a key objective, this is not necessarily surprising.

²⁸ For a thorough understanding of each theoretical perspective, see: Grych, J.H., & Fincham, F.D. (1990). Marital conflict and children's adjustment: A cognitive-contextual framework. *Psychological Bulletin*, 108(2), 267–290; Davies, P.T., & Cummings, E.M. (1994). Marital conflict and child adjustment: An emotional security hypothesis. *Psychological Bulletin*, 116(3), 387–411; Harold, G.T., & Conger, R.D. (1997). Marital conflict and adolescent distress: The role of adolescent awareness. *Child Development*, 68(2), 333–350.

²⁹ Harold et al. (2016). Op. cit.

³⁰ Ibid

³¹ Ibid

The limited evidence of programmes demonstrating impact on child outcomes may be because:

- relationship support services are rarely commissioned with the explicit aim of improving child outcomes; consequently, capturing these outcomes is not typically a focus of their work nor is it included in key performance indicators
- service providers find it difficult to capture the 'secondary' impact of their work, particularly when considering the challenge of capturing outcomes for children which are likely to be longer term and not immediately evident
- measuring outcomes for children is generally perceived to be challenging, particularly in the case of children who are not directly involved in the intervention or young children who cannot articulate their feedback
- limited funding means that programme providers do not have sufficient resources to embed validated measures of child outcomes within their evaluations.

Although these reasons are justifiable, given how parental conflict can impact children via the Family Stress Model (see figure 2.1), **our view is that it is important to specifically measure child outcomes.** If your area is delivering a programme to reduce parental conflict, but you only measure the quality of the interparental relationship and the level of conflict between parents, you won't know if your service has benefited children unless you test this directly – for example, by measuring children's emotional and behavioural problems.

2.3 Outcomes we are focusing on in this guide

A key conclusion of the 2016 What Works review is that children's wellbeing could be greatly improved through interventions that target the interparental relationship.³² For this reason, we have decided to focus this guide on how to measure changes in children's emotional and behavioural problems, and in particular on their **internalising and externalising behaviours** (for a detailed description of these, see the list in section 2.1).

As noted above, there are several other child outcomes affected by parental conflict.³³ However, we have chosen to focus on those outcomes which appear to be most commonly affected by children who witness destructive and ongoing parental conflict. We have also chosen to focus on more proximal (immediate) outcomes, rather than distal (longer-term) outcomes such as academic failure, substance misuse and mental health difficulties in later life. Importantly, the measures we are focusing on in this guide are aimed at a non-clinical population, and the guide does not aim to assess measures for children with learning disabilities, neurodevelopmental disorders, and physical or mental conditions diagnosed by a clinician.

Alongside child outcomes it is also important to measure interparental relationship outcomes, as improving these tends to be the primary outcome of interventions seeking to reduce parental conflict. For the purposes of this guide, the outcomes we have decided to focus on are informed by the Family Stress Model, which shows that the quality of the interparental relationship, including parental conflict, can impact on parenting to affect children's outcomes. In order to keep our scope narrow, rather than focusing on parenting as a whole, we have opted to concentrate on co-parenting practices only.

Consequently, in scope for this guide were measures assessing the **interparental relationship quality** (such as satisfaction, commitment, communication, respect, and the extent of agreements and disagreements between individuals); **parental conflict** (specifically, the freq-

³² Ibid

³³ Ibid

uency and intensity of conflict, conflict resolution skills, and the children's perception of and adjustment to parental conflict); and **co-parenting practices** (for example, the parents' ability to cooperate and communicate when caring for their children). Nonetheless, it is worth bearing in mind that many of the measures included in this guide have items assessing more than one of these three outcomes. However, to facilitate the selection of appropriate measures, we have grouped measures according to the outcome most prevalent within each measure.

The following are outcomes we are not focusing on within this guide, but which would be relevant to measure if you had the necessary time and resources (and if they were the intended outcomes of the interventions you were evaluating).

Child outcomes

- social and interpersonal relationship skills (reflected through the quality of peer relationships, for example)
- academic performance (measured, for example, through exam results and teacher reports)
- physical health (including rates of illness, sleep patterns and fatigue, as well as risky health behaviours such as smoking, substance misuse and early sexual activity).

Interparental (or couple) outcomes

- domestic abuse (including destructive parental conflict that has become violent or abusive).

Parent outcomes

- mental health and psychological wellbeing (for example, parental stress and anxiety)
- parenting practices (such as harsh or aggressive parenting)
- parenting self-efficacy (including parent's belief in their ability to parent successfully).

Family outcomes

- family functioning (such as the quality of parent–child interactions and sibling relationships)
- home environment (including family stress and the provision of stimulating learning experiences).

For more information on measures assessing outcomes not included in this guide, we have compiled a list of potentially helpful resources you should consider engaging with (see section 5.4).

In summary, the outcomes we are focusing on in this guide, and which we have based our selection of measures on, include:

Child outcomes

- Internalising behaviours
- Externalising behaviours

Interparental outcomes

- Relationship quality
- Parental conflict
- Co-parenting practices.

3. Why measure?

There are various reasons why you may be interested in using measurement tools. In the context of interparental relationships, for example, measures could be used for:

- **observational** purposes, such as to provide practitioners with a better understanding of how couples/co-parents communicate with one another
- **population surveillance** purposes, such as to understand population trends in terms of the frequency and intensity of conflict between couples/co-parents so that local areas are better able to meet the needs of their population
- **screening or diagnostic** purposes, such as to identify whether couples are in domestic abusive relationships
- **eligibility** purposes, such as to determine whether couples/co-parents may benefit from relationship support and if so, what type of support would best suit them
- **therapeutic** purposes, such as to provide practitioners with insight into the couple/co-parenting relationship so that they are better able to tailor the support provided
- **monitoring** purposes, such as to allow both the practitioner and participant to assess and reflect on how they are progressing in time
- **evaluation** purposes, such as to measure change over time and assess the impact of interventions
- **feedback** purposes, such as to collect information from participants on their experience of the intervention, including the practitioner delivering it.

Some measures may be suitable for a range of purposes; however, it is your responsibility to be clear on what your measurement purpose is so that you are better able to select the most appropriate tool.

Carrying out an initial risk assessment

Many parents only seek interparental relationship support when they experience a crisis or are on the brink of (or indeed going through) separation or divorce. For many, this is a critical transition point and a stressful period of time when they may be particularly vulnerable to mental health problems, and at a heightened risk of suicidal ideation, substance misuse, domestic abuse and self-harm. Therefore, it is vital that an initial risk assessment is carried out prior to enrolment in an intervention or service, to ensure the individual's suitability to take part.

Although it is beyond the scope of this report to review risk assessment methods, it is a matter of good practice to have well-established procedures for the assessment and handling of high-risk clients.

For the purposes of this guide we are **interested in measurement for evaluation purposes, to assess the progress of individual families taking part in programmes targeting the interparental relationship**. The reasons why you might be particularly interested in measuring whether outcomes have changed in response to an intervention or service may be because:

- you want to **know** that the support you provide is effective
- you want to be able to **prove** to others that your intervention or service provides benefits for families that are meaningful and therefore worth continued investment.

Measuring how families are affected by local services is essential to determine if they are working for the people they are designed to help. Furthermore, given the link between parental conflict and child outcomes, it is important to ensure that you measure changes brought by the intervention in both parents and children. Crucially, this involves selecting a measure that is consistent with the intervention you are delivering and the population you are working with, and using this tool to measure if the outcomes of the families involved have improved after the intervention. For simplicity, these measures are often referred to as **outcome measures**.

With the design and delivery of new interventions and relationship support services, including the rollout of the national Reducing Parental Conflict Programme, **there is a crucial opportunity to build the evidence base in the UK on what works to improve outcomes for children by reducing parental conflict**. While you may be fairly sure your service helps, without measuring outcomes robustly as part of a strong evaluation, you won't know if or why your services worked as intended. This means using validated measures to monitor the progress of individual children and their parents, both before and after their family has used your service. This can be done by frontline professionals delivering your service or by external evaluators you commission.

3.1 What are validated measures and why is using them important?

There is a growing number of measurement tools that assess a whole host of outcomes. Many of these tools, however, have not yet been adequately validated. For a measure to be validated, it needs to have been carefully and independently subjected to a series of statistical tests, verifying that it is both valid and reliable. Only when measures have gone through this research process can we say anything about their 'psychometric features'. Until then, we cannot be sure that the results a measure yields are a valid or reliable representation of the outcome being assessed. The quality of the information provided by a measure therefore depends, at least partly, on its psychometric features.

When selecting measures to use in your impact evaluation – be it a pre/post study or a randomised controlled trial – it is thus crucial that you **choose measures which have had their psychometric features validated**. This is because without valid and reliable measures it is not possible to have confidence in the findings and conclusions of studies. For the purposes of this guide, we were interested in four key psychometric features, which are the ones used in most of the assessment criteria adopted by other clearinghouses.

- **Internal consistency:** the degree to which items designed to measure the same outcome relate to one another. If a measure has several items designed to assess symptoms of anxiety, then all these items should be sufficiently correlated to one another so that we can confidently say that the tool measures symptoms of anxiety in a consistent manner. If measurement items are not sufficiently correlated with one another, then issues of inconsistency arise, with responses to different items yielding incongruent results which undermine the reliability of the data collected. In some cases, measurement tools are comprised of several subscales assessing multiple outcomes.

A measure of parental conflict assessing both the frequency and intensity of conflict might, for example, be comprised of two subscales – the frequency and intensity subscales. In this instance, the feature of internal consistency would need to be tested for each subscale independently.

- **Test-retest reliability:** the extent to which the outcome of a measure is stable over time, assuming no important intervening event has occurred. If a measure is able to produce the same result when repeated multiple times under the same or similar circumstances, we deem it to have established test-retest reliability (also referred to as stability). For example, a reliable measure of co-parenting should elicit the same (if not identical) result if it was to be administered in a month's time, unless something significant (like an intervention) changed the participant's co-parenting practices. Measurement tools must therefore have established test-retest reliability to provide a reliable estimate of change over time. Without this, it would be very difficult to establish whether any identified change was due to the unreliability of the measure or the effectiveness of the intervention.

A note on reliability

Internal consistency and test-retest reliability are different psychometric features of a measure's reliability, which refers to how stable, consistent or reproducible it is. Unreliable measures therefore do not accurately capture change because they are unstable and vary randomly over time.

Another psychometric feature of a measure's reliability is its interrater reliability, or the extent to which different observers making assessments on an individual come to the same or similar conclusions. Although an important psychometric feature, we did not assess within this guide measures for their interrater reliability since we did not select observational measures where this would be an important feature to examine.

- **Validity:** the extent to which a measure describes or quantifies what it claims to measure. Knowing that your chosen measure is accurately quantifying the outcomes it claims to assess is crucial. A measure that claims to assess how anxious or depressed a child is feeling, but instead only asks children how often they take part in sport and extracurricular activities, for example, is not a valid measure of anxiety and depression. Invalid measures do not measure what they claim to measure and so are not true measures of the outcomes you are interested in. When considering validity, it is important to note that there are several different types of validity, including criterion validity and construct validity.³⁴
 - **Criterion validity** is the extent to which the score of a given measure is related to some external criterion – usually a 'gold standard' assessment such as a clinical observation. Assessing criterion validity is a challenge, mainly because it is costly and requires a 'gold standard' measure which is not always available. For this reason, not all measures have criterion validity, meaning that they have not been verified against clinical observations of actual behaviours.
 - **Construct validity** should be used to provide evidence of validity in situations where there is no 'gold standard'. In producing this guide, we mainly came across tests of construct validity, which refers to whether a measure provides the expected scores based on existing knowledge about the outcome under assessment. It is therefore frequently defined by the degree to which the scores of a measure are consistent with

³⁴ Souza, A.C.D., Alexandre, N.M.C., & Guirardello, E.D.B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, 26(3), 649–659.

hypotheses, for instance with regard to relationships to scores of other instruments, or differences between relevant groups.³⁵

- » One type of construct validity (and the one we most frequently encountered) is **convergent validity**, which refers to the extent to which the scores on a particular measure correlate with scores on another measure in a manner that is consistent with theoretically derived hypotheses concerning the measured outcomes. For example, two different measures of anxiety should correlate well with each other, so long as they are assessing the same (or similar) type of anxiety problem.
- » Another type of construct validity is **discriminant validity**, which refers to the extent to which a measure can discriminate between people who are expected to score high on a measure from those who are expected to score low. For example, people who have a diagnosis of anxiety disorder should score higher (on average) on a measure of anxiety than those who do not have such a diagnosis.
- **Sensitivity to change: the degree to which a measure is able to detect change within the timeframe of the programme under evaluation.** This feature is largely dependent on the reliability and validity of a measure, as assessed using the psychometric features described above. There are, however, some factors which contribute to a measure's sensitivity. Longer measures with more items, for example, are more reliable and sensitive to change over short periods of time. There are of course other issues with long measures such as the burden they inflict on participants. For the purposes of this guide, we have focused on identifying whether the selected measures were able to detect significant changes over time in impact evaluations of short-term interventions, since the interventions being trialled as part of the Reducing Parental Conflict Programme are relatively short, with the longest lasting 16 weeks.³⁶

Ultimately, the reason we encourage the use of validated measures is because **unreliable and invalid measures will yield unclear results, and it will not be possible to determine whether any apparent improvement in outcomes is due to a true effect of the intervention under evaluation, a product of measurement error, or bias.** The selection and appropriate use of validated measures is therefore a core pillar of any robust evaluation because without valid and reliable measures it is not possible to have confidence in the findings and conclusions of studies.

How do EIF's evidence standards take measurement tools into consideration?

For a study to be considered as providing at least preliminary evidence (that is, an EIF evidence rating of level 2 or higher), at least one significant child outcome must be identified on a measure which is valid, reliable and appropriate for the anticipated outcomes and population. Administrative data (for example, school absences or exam results) does not need to have established validity and reliability, but it is important that the data are described in detail and their sources well defined. In very specific circumstances EIF holds to less strict standards of validation measures which assess phenomena that are directly observable or recollectable, such as self-reports of substance use. For more details, see the EIF Guidebook.³⁷

³⁵ Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., & de Vet, H.C.W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63, 737–745.

³⁶ For more information on the eight interventions selected by DWP to be trialled in the UK as part of the national Reducing Parental Conflict Programme, see: <https://www.eif.org.uk/files/pdf/cg-rpc-4-3-face-to-face-support-interventions.pdf>

³⁷ For more details on the EIF evidence standards, see: <https://guidebook.eif.org.uk/eif-evidence-standards>

3.2 Additional features of desirable measures

While the most important feature of any measurement tool is that it is both valid and reliable, there are several other more practical features that are important to consider when deciding what measures to use.

In addition to the psychometric features listed above, we assessed the measures included in this guide against a number of implementation features which we felt were important for local areas to consider from a practical perspective.

- **Brevity: the time taken to complete a measure.** Brevity is related to the number of items in a measure and the response options available for each item. Excessively long measures can lead to measurement fatigue, which can cause some respondents to not complete all the items or questions, negatively impacting on the quality of the data. Less burdensome measures such as the abbreviated version of the Dyadic Adjustment Scale (DAS-7), which assesses relationship quality and includes only seven items, can be completed quickly and is therefore quite helpful from a practical standpoint. Nonetheless, it is important to bear in mind that brevity must be balanced against the validity and reliability of a measure, as longer measures tend to have higher psychometric reliability, and to be more sensitive to change.
- **Availability: the extent to which a measure is freely available.** Measures that are freely available are appealing, but not necessarily better than those only available at a cost. Measures that are not freely available have different charging structures, with some having a 'one off' or 'per use' fee. Sometimes the fee is dependent on the format of the measure (for example, if it is completed by paper or online) or on the motive for use (such as whether it will be used for research or clinical purposes).
- **Ease of scoring: the extent to which a measure is easy to score and interpret.** Measures that are easy to score tend to involve simple additions and subtractions. However, there are some measures which require complex calculations or scoring by a professional (for example, a researcher or practitioner) who has undergone specific training. This can increase the precision and sensitivity of the measure, but also creates an additional burden in the scoring and associated cost of using that measure.
- **Used in the UK: the extent to which a measure has been used in the UK context.** Since the aim of this guide is to identify and assess some of the measures that could be used by UK local areas to evaluate the impact of programmes targeting the interparental relationship, we felt it was important to determine whether the measures included in this guide had been previously used in the UK context, with a representative sample population. Importantly though, this feature does not refer to whether a measure *could be used* in the UK, just if there is evidence to suggest that it *has been used*.

For details on how we arrived at these psychometric and implementation features, what criteria we used to define each feature, and how we assessed the measures included in this guide against the criteria, please see the methodology in appendix C.

4. Measures

There is an overwhelming number of measures that can be used to assess child and couple outcomes. **We have assessed some of the most relevant and frequently used measures that could be used to assess progress for individual families taking part in programmes targeting the interparental relationship.** In this chapter of the guide, we provide you with a short description of how we selected the measures included here, as well as how we assessed them to establish if they are well validated (both valid and reliable) and practical to implement. For a more detailed description of our selection and assessment processes, please see the methodology in appendix C.

4.1 Selecting relevant measures

To identify a longlist of suitable measures, we collated measures that were:

- used in impact evaluations of programmes targeting interparental relationships, which were identified in our 2017 review on parental conflict in the context of poverty,³⁸ but mainly drawn from our 2016 review in collaboration with Professor Gordon Harold³⁹
- used in evaluations of the eight face-to-face interventions currently being trialled in the UK as part of DWP's Reducing Parental Conflict Programme⁴⁰
- used in DWP's own evaluation of the national programme⁴¹
- suggested by our advisory members, comprised of subject-matter experts and local authority leaders.

In total, we identified 233 measures. This longlist of measures was then tested against a set of inclusion criteria to arrive at a more manageable number for assessment. Where possible, we prioritised measures known to be commonly used by local areas and representing a broad range of outcomes and target populations, to consider both intact and separated couples as well as children of different age groups. For details of our inclusion criteria, please see appendix C.

Selected measures

Ultimately, we included 18 measures, seven assessing child outcomes and 11 assessing interparental (or couple) relationship outcomes (table 4.1). In our view, the final selection of measures included in this guide encapsulates some of the most relevant and frequently used tools. However, **it is not an exhaustive list of all the measures you could consider.** It is also important to note that we selected these measures based on our predefined outcomes of interest (see appendix C), and so there may be measures assessing other outcomes of importance that we have not considered here (for example, parental mental health). Due to the methods used to identify our longlist of measures and arrive at a final number for inclusion here, **it is also possible we have missed out important measures that assess our predefined outcomes of interest.** For example, we have not included measures referring to children under the age of 2 years for whom there is now evidence that parental conflict can affect.

³⁸ Acquah et al. (2017). Op. cit.

³⁹ Harold et al. (2016). Op. cit.

⁴⁰ See: <https://www.eif.org.uk/files/pdf/cg-rpc-4-3-face-to-face-support-interventions.pdf>

⁴¹ See box 'Measures used in DWP's evaluation of the national Reducing Parental Conflict Programme', later in this section.

TABLE 4.1

Measures selected for inclusion in this guide

| Measures | Total |
|--|-----------|
| Child outcome measures | 7 |
| Internalising and externalising behaviours | 2 |
| Child Behaviour Checklist (CBCL/6–18) | |
| Strengths and Difficulties Questionnaire (SDQ) | |
| Internalising behaviours | 3 |
| Mood and Feelings Questionnaire (MFQ) | |
| Revised Child Anxiety and Depression Scale (RCADS) | |
| Screen for Child Anxiety Related Emotional Disorders (SCARED) | |
| Externalising behaviours | 2 |
| Eyberg Child Behavior Inventory (ECBI) | |
| Parent Daily Report (PDR) | |
| Interparental (or couple) relationship outcome measures | 11 |
| Relationship quality | 6 |
| Couples Satisfaction Index (CSI-16) | |
| Dyadic Adjustment Scale (DAS-32) | |
| Dyadic Adjustment Scale (DAS-7) | |
| Golombok Rust Inventory of Marital State (GRIMS) | |
| Relationship Quality Index (RQI) | |
| Marital Adjustment Test (MAT) | |
| Parental conflict | 2 |
| Children's Perception of Interparental Conflict Scale (CPIC) | |
| O'Leary Porter Scale (OPS) | |
| Co-parenting | 3 |
| Parenting Alliance Measure (PAM) | |
| Parent Problem Checklist (PPC) | |
| Quality of Co-parental Communication Scale (QCCS) | |

Child outcome measures

The child outcome measures included in this review have been designed to assess **internalising and/or externalising behaviours**, which are key outcomes affected by destructive and ongoing parental conflict. While some of the included measures assess only internalising or externalising behaviours (such as MFQ and ECBI), some assess both (such as CBCL/6–18 and SDQ). Many of the child outcome measures included in this guide also have a range of available versions that can be used for children of different ages and with different respondents (parents, teachers or the children themselves). Due to time and resources constraints, **we have only assessed one version of each selected measure** – either the parent report or the child self-report version. However, in the summary reports (presented in appendix A) we have listed all the available versions identified from our review of the literature. When reviewing the summary reports, **please be aware that our assessments are not transferable to other versions of the same measure not assessed in this guide**. It should therefore not be assumed that other versions of the same measure would receive the same rating. For example, although our assessment of the psychometric features of the CBCL/6–18 were very positive with all features meeting our criteria, the same might not be true for the teacher-rated version of this measure (Teacher's Report Form, or TRF). A high rating of one version does not imply that other versions of the same measure are equally valid and reliable.

Interparental (or couple) relationship outcome measures

A variety of interparental (or couple) relationship outcome measures have been developed for use across different contexts, with measures ranging from short instruments asking respondents to rate how satisfied they are in their relationship (for instance CSI-16), to more complex measures comprising several subscales and assessing a variety of factors. Of all the interparental measures included in this guide, only one (the CPIC) is completed by the child.

In this review we have included a selection of measures examining the **interparental relationship quality, parental conflict** and **co-parenting practices**. It is worth bearing in mind that many of the selected measures have items assessing more than one of the three outcomes (interparental relationship quality, parental conflict and co-parenting practices), as reported in table 4.3. However, to facilitate the selection of appropriate measures, we have grouped measures according to the most prevalent outcome. While it has not been possible to include all measures assessing these outcomes, the measures selected are some of the most commonly used. As with the child outcome measures, **we have only assessed one version of each selected measure and so it should not be assumed that other versions of the same measure would yield an equal rating**. The only exception to this is the DAS, where we have assessed both the short DAS-7 and the original DAS-32. We assessed both these versions because we know that the DAS is one of the most commonly used relationship measures, and wanted to ensure we captured a long and shortened version in order to address different measurement needs. We are also aware that both these versions are being used as part of DWP's evaluation of the national programme.

Measures used in DWP's evaluation of the national Reducing Parental Conflict Programme

As part of the national evaluation of the Reducing Parental Conflict Programme, DWP are using the interparental and child outcome measures included below. Measures marked with an asterisk have been assessed by EIF and included in this guide. The other measures have not been included as we did not find sufficient publicly available information for these measures to be in scope of this review, based on our assessment criteria.

For intact couples:

- Dyadic Adjustment Scale (DAS-7)*
- Satisfaction Subscale of the Dyadic Adjustment Scale (DAS-32)*
- O'Leary Porter Scale (OPS)*
- Warmth and Hostility Subscales of the Iowa Family Interaction Rating Scale (IFIRS)

For separated couples:

- Quality of the Coparental Communication Scale (QCCS)*
- Co-parenting Communication
- Frequency and Breadth of Conflict Scale

For children:

- Strengths and Difficulties Questionnaire (SDQ)*

4.2 Assessing selected measures against our psychometric and implementation criteria

Once we had selected the measures to include in this guide, we assessed each against our predefined criteria to examine whether they were valid, reliable and practical to implement.

Our measurement assessment criteria (see appendix C, table C.2) focused on the key psychometric and implementation features described in chapter 3. They were developed by taking into account the criteria used by other clearinghouses, and discussion with a group of psychometric experts, to create a set of demanding but achievable cut-off values which measures had to reach in order to be awarded the criterion. Our psychometric experts (see appendix D) also provided guidance on issues regarding the application of the criteria and ensuring that we were applying it consistently to all measures.

When applying our assessment criteria to each selected measure, we first developed a search strategy to ensure that we identified the most relevant studies describing the design, development and psychometric testing of the selected version of each measure. Once the relevant studies had been identified, we extracted pertinent information concerning the psychometric and implementation features of the selected version of each measure, and rated each criterion item according to the following rating system:

| | |
|---|--|
| ✓ | We found clear evidence that the selected version of the measure had met the criterion |
| ✗ | We were unable to find evidence showing that the selected version of the measure had met the criterion |
| ? | We found insufficient evidence to determine if the criterion had been met |

Although we sought to select measures commonly used by those commissioning or delivering relationship support services, that does not necessarily mean that all included measures have passed our rigorous criteria. Indeed, **there are several measures with psychometric or implementation features which have not passed our thresholds, or where we were unable to find sufficient evidence to firmly establish if the criteria had been met.** It is also important to note that the methodological approach used in this rapid review did not involve an exhaustive search of the literature. **There is therefore a risk that we have missed relevant evidence when applying our assessment criteria against the selected measures.** For more details on how we developed our criteria and search strategy, please see the methodology in appendix C.

4.3 A snapshot of the validity, reliability and practicality of our selected measures

If you are interested in the outcomes we have focused on within this guide and are considering selecting some of the measures included here, we have tried to make your life simpler by including **summary tables** of our independent assessments of the 18 outcome measures. Table 4.2 presents the child outcome measures and table 4.3 shows the interparental outcome measures. Hopefully you will find that these two tables allow for easy comparison between measures, and help you make sense of the detailed information we collated.

To help you navigate and digest this information more easily, we have also created two **decision support tools**, one for the child and another for the interparental measures (see figure 4.1 and 4.2, respectively). It is important to remember, however, that these tools are mainly designed to guide your thinking, and that only validated measures which are consistent with the intervention's theory of change and target population should be selected (for more on that, see chapter 5).

The decision support tools have therefore been designed solely based on the implementation features of the measures assessed, which hopefully will help you interpret the information through a different lens, considering what might be most practical and useful for your context. Measures with equal psychometric scores assessed against our criteria, are not necessarily equally applicable in all contexts. A reading of table 4.2, for example, would imply there is no reason to use the CBCL/6–18 over the SDQ as both get full scores on psychometric features, but the SDQ rates more highly on implementation features as it is quick to administer, freely available and easy to score. The CBCL/6–18, however, is a very detailed measure with over 100 items and so it provides a more in-depth assessment, including more severe psychiatric symptoms such as compulsions and hallucinations. The CBCL/6–18 is thus more likely to be suitable in situations requiring a more detailed assessment of a more comprehensive range of problems.

What is a theory of change or logic model?

It is not uncommon for people to use the terms 'theory of change' and 'logic model' interchangeably. This is because both terms explain processes that create a theoretical link between an intervention's activities and its intended short- and long-term outcomes. A logic model is a statement of what a programme or service consists of (in terms of inputs, activities and outputs) and what a programme intends to achieve (short- and long-term outcomes), while a theory of change describes how and why a programme is expected to achieve its desired effects. For more information on theories of change and logic models, see Step 1 and Step 2 of the *10 steps for evaluation success* guidance report.⁴²

Throughout this guide, the selected measures are presented always in the same order. Child outcome measures are presented first, grouped by outcome (internalising and externalising behaviours; internalising behaviours; and externalising behaviours) and within that, in alphabetical order. Next we present the selected interparental (or couple) relationship outcome measures, again grouped by outcome (relationship quality; parental conflict; and co-parenting) and alphabetical order.

More information on each measure can be found within the individual **summary reports**, included in appendix A. The **full measurement reports**, including details of the evidence used to assess each measure against our predefined psychometric and implementation criteria, are included in attachment to this report and can be downloaded from the EIF website.⁴³

We strongly encourage you to engage with this more detailed evidence so that you understand the strengths and limitations of the measures you are considering, and can make more informed and evidence-based decisions.

⁴² See: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

⁴³ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

FIGURE 4.1

Decision support tool for child outcome measures

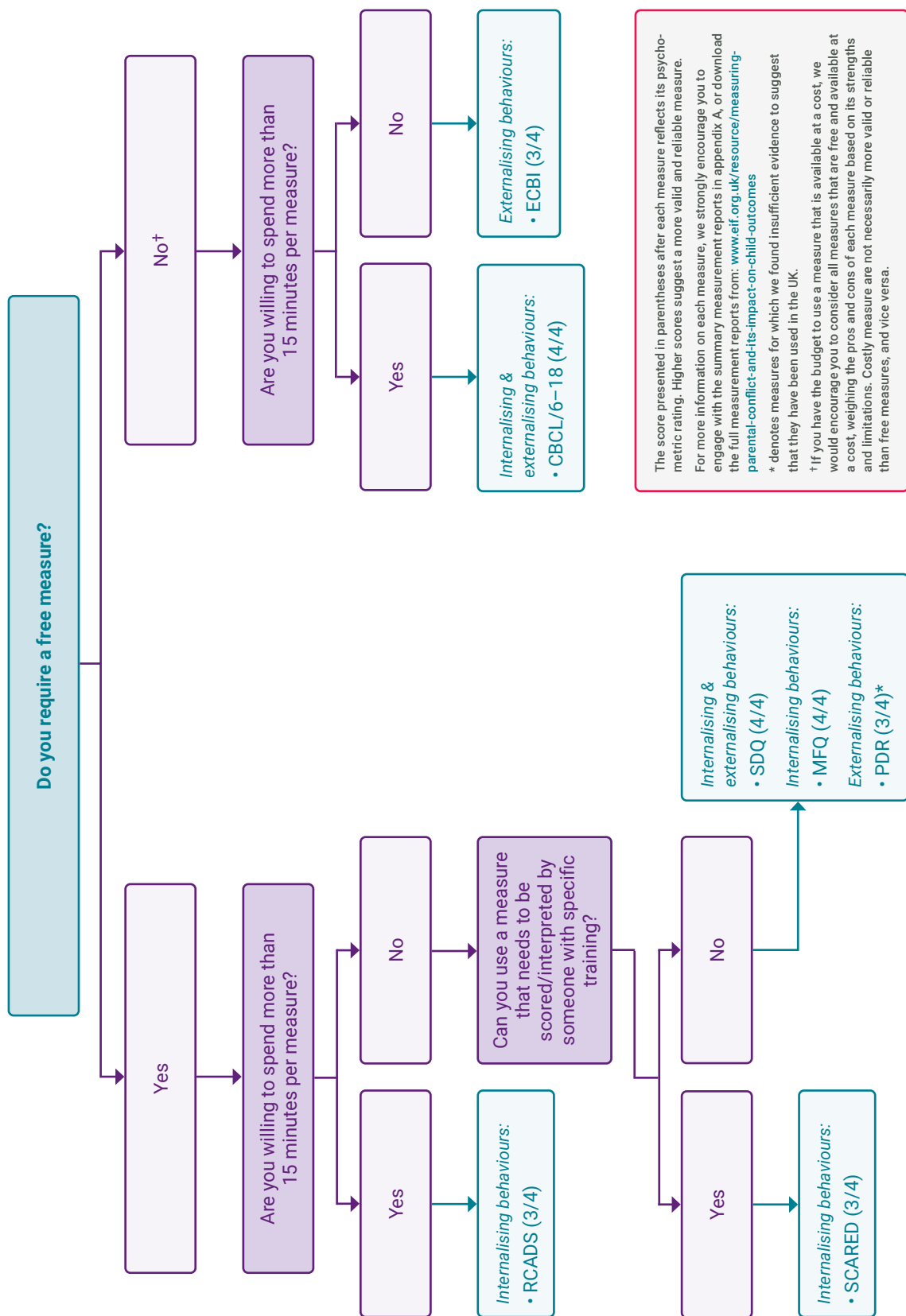


FIGURE 4.2

Decision support tool for interparental relationship outcome measures

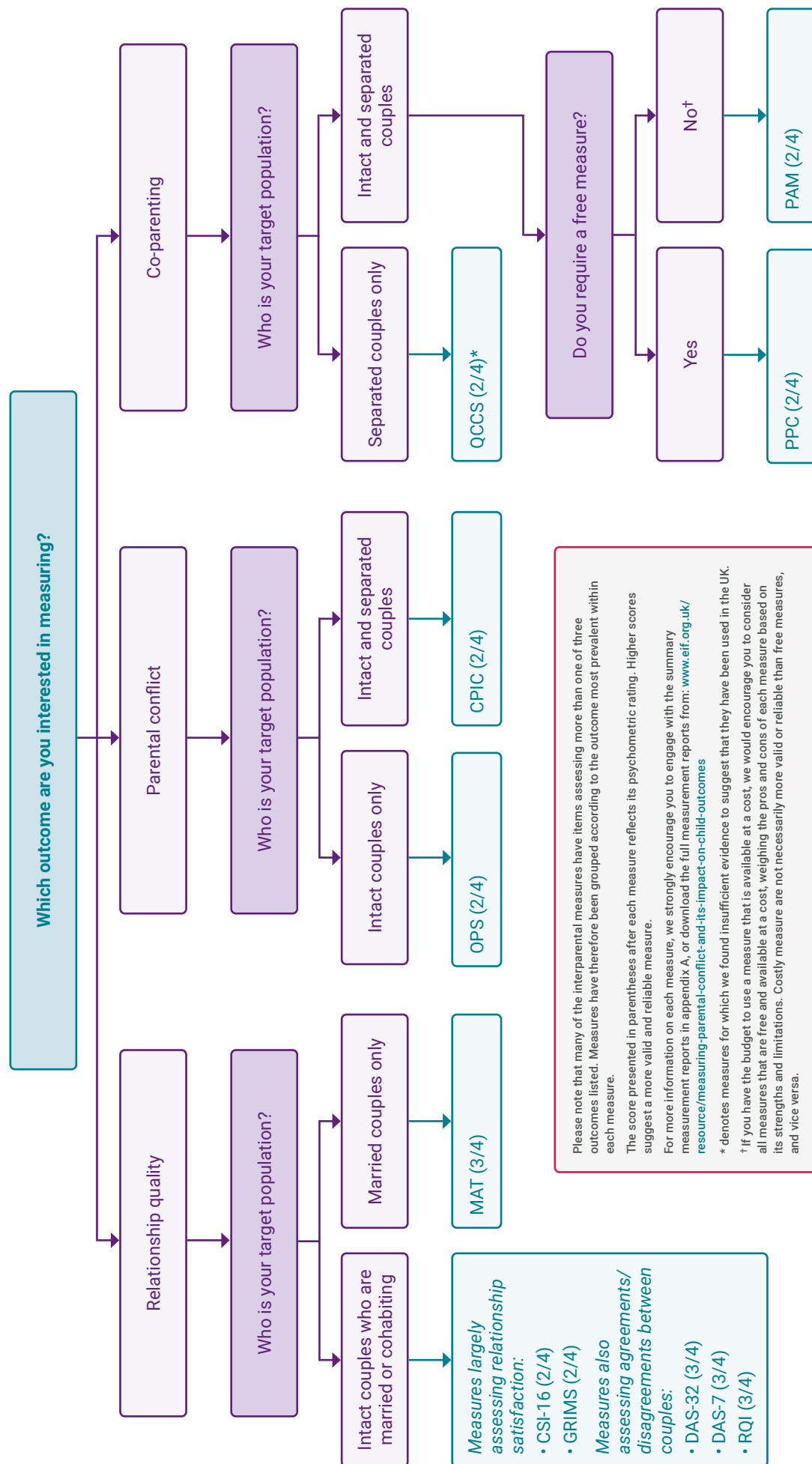


TABLE 4.2

Child outcome measures

| | | Psychometric features How valid and reliable is the measure? | | | | | Implementation features How practical is the measure? | | | |
|--|---|--|--|--|--|------------------------------------|--|------------------------------------|-----------------------------|----------------|
| | | Internal consistency (scale) | Internal consistency (subscales) | Test-retest reliability | Validity | Sensitivity to change | Brevity | Availability | Ease of scoring | Used in the UK |
| | | Do the items designed to measure the same outcome relate to one another? | Are the outcomes stable over time? | Does it measure what it claims to measure? | Can it detect important changes over time? | Does it take less than 15 minutes? | Is it freely available? | Is it easy to score and interpret? | Has it been used in the UK? | |
| Internalising & externalising behaviours | Child measures Outcomes assessed CLICK ON A MEASURE FOR MORE DETAIL | Respondent: Who should complete this version of the measure? | Target population: Who is the measure developed for? | | | | | | | |
| | Child Behaviour Checklist (CBCL/6-18) This measure assesses the child's emotional and behavioural problems, including (but not limited to) anxiety, depression, hyperactivity and aggressive behaviour. | Parents of children aged 6-18 years | Children aged 6-18 years | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ |
| | Strengths and Difficulties Questionnaire (SDQ) This measure assesses the child's emotional and behavioural problems, including (but not limited to) prosocial behaviour, hyperactivity, conduct and peer problems. | Parents of children aged 4-17 years | Children aged 4-17 years | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

| | | | Psychometric features How valid and reliable is the measure? | | | | Implementation features How practical is the measure? | | | | |
|--------------------------|--|--|--|------------------------------------|--|--|--|-------------------------|------------------------------------|-----------------------------|----------------|
| | | | Internal consistency (scale) | Internal consistency (subscales) | Test-retest reliability | Validity | Sensitivity to change | Brevity | Availability | Ease of scoring | Used in the UK |
| | | | Do the items designed to measure the same outcome relate to one another? | Are the outcomes stable over time? | Does it measure what it claims to measure? | Can it detect important changes over time? | Does it take less than 15 minutes? | Is it freely available? | Is it easy to score and interpret? | Has it been used in the UK? | |
| Internalising behaviours | Child measures Outcomes assessed CLICK ON A MEASURE FOR MORE DETAIL | Respondent: Who should complete this version of the measure? | Target population: Who is the measure developed for? | Children aged 8-18 years | Children aged 8-18 years | N/A | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | Mood and Feelings Questionnaire (MFQ) This measure assesses the child's feelings, thoughts and behaviours associated with depression. | Children aged 8-18 years | Children aged 8-18 years | ? | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ |
| | | Revised Child Anxiety and Depression Scale (RCADS) This measure assesses the child's symptoms of depressive and anxiety disorders. | Children aged 8-18 years | Children aged 8-18 years | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | Screen for Child Anxiety Related Emotional Disorders (SCARED) This measure assesses the child's symptoms of different anxiety disorders. | Children aged 8-18 years | Children aged 8-18 years | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ |
| Externalising behaviours | | Eyberg Child Behavior Inventory (ECBI) This measure assesses the frequency and severity of disruptive behaviours manifested by the child. | Parents of children aged 2-16 years | Children aged 2-16 years | ✓ | ✓ | ? | ✓ | ✗ | ✗ | ✓ |
| | | Parent Daily Report (PDR) This measure assesses negative behaviours displayed by the child at home. | Parents of children aged 4-10 years | Children aged 4-10 years | N/A | ✓ | ? | ✓ | ✓ | ✓ | ? |

TABLE 4.3

Interparental (or couple) relationship outcome measures

| | | | Psychometric features How valid and reliable is the measure? | | | | Implementation features How practical is the measure? | | | | | |
|--|--|--|--|--|--|------------------------------------|--|--|------------------------------------|-------------------------|------------------------------------|-----------------------------|
| | | | Internal consistency (scale) | Internal consistency (subscales) | Test-retest reliability | Validity | Sensitivity to change | Brevity | Availability | Ease of scoring | Used in the UK | |
| Interparental measures Outcomes assessed CLICK ON A MEASURE FOR MORE DETAIL | | | Respondent: Who should complete this version of the measure? | Target population: Who is the measure developed for? | Do the items designed to measure the same outcome relate to one another? | Are the outcomes stable over time? | Does it measure what it claims to measure? | Can it detect important changes over time? | Does it take less than 15 minutes? | Is it freely available? | Is it easy to score and interpret? | Has it been used in the UK? |
| Relationship quality | | | Adults in a relationship | Intact couples | N/A | ? | ✓ | ? | ✓ | ✓ | ✓ | ✓ |
| | | | Adults in a relationship | Intact couples | ✓ | ✓ | ✓ | ? | ✓ | ✓ | ✓ | ✓ |
| | | | Adults in a relationship | Intact couples | N/A | ? | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | Adults in a relationship | Intact couples | N/A | ? | ? | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | Adults who are married | Married couples | N/A | ✓ | ? | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | | Adults in a relationship | Intact couples | N/A | ? | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ |

| | | Psychometric features How valid and reliable is the measure? | | | Implementation features How practical is the measure? | | | | | | |
|---|---|---|---|--|--|--|--|------------------------------------|-------------------------|------------------------------------|-----------------------------|
| | | Internal consistency (scale) | Internal consistency (subscales) | Test-retest reliability | Validity | Sensitivity to change | Brevity | Availability | Ease of scoring | Used in the UK | |
| Interparental measures Outcomes assessed CLICK ON A MEASURE FOR MORE DETAIL | | Respondent: Who should complete this version of the measure? | Target population: Who is the measure developed for? | Do the items designed to measure the same outcome relate to one another? | Are the outcomes stable over time? | Does it measure what it claims to measure? | Can it detect important changes over time? | Does it take less than 15 minutes? | Is it freely available? | Is it easy to score and interpret? | Has it been used in the UK? |
| Parental conflict | Children's Perception of Interparental Conflict Scale (CPIC) Parental conflict from the child's point of view, particularly frequency, intensity, resolution and perceived threat. Also how the child responds to the conflict, including questions around self-blame and coping mechanisms. | Children aged 9-17 years with intact or separated parents | Intact and separated couples with children | ✓ | ✓ | ? | ? | ✗ | ✓ | ✓ | ✓ |
| | O'Leary Porter Scale (OPS) The frequency of couple hostility observed by the child, including quarrels, sarcasm and physical abuse. | Parents in a relationship | Intact couples with children | ✓ | N/A | ? | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Parenting Alliance Measure (PAM) How cooperative, communicative and mutually respectful parents are when caring for their children. | Parents who are in a relationship or are separated | Intact and separated couples with children | ✓ | ✓ | ? | ✓ | ✓ | ✗ | ✓ | ✓ |
| | Parent Problem Checklist (PPC) The extent of agreement or disagreement between the parents over child-rearing issues. | Parents who are in a relationship or are separated | Intact and separated couples with children | ✓ | N/A | ? | ✓ | ✓ | ✓ | ✓ | ✓ |
| Co-parenting | Quality of Co-parental Communication Scale (QCCS) The extent of mutual support and hostility over child-rearing issues with the former spouse. | Parents who are separated | Separated couples with children | ✓ | ? | ✓ | ? | ✓ | ✓ | ✓ | ? |

Although not formally included in this guide, the Family Star (Relationships) is one of the suite of Outcome Stars developed by Triangle Consulting and commonly used by local areas working with families. This is a relatively new measure, published in September 2018, and is mainly focused on assessing how well parents who are in an intact or separated relationship manage their feelings and any conflicts that arise around child-rearing issues. Currently, there is very limited information on the psychometric features of this tool and therefore we did not feel it made sense for us to assess it as part of this guide. Nonetheless, below we have collated some information on the Family Star (Relationships), which we hope will be useful to local areas.

Family Star (Relationships)

The Family Star (Relationships) was developed by Triangle in collaboration with the Essex and Hertfordshire County Councils, as part of the Local Family Offer work funded by the DWP. To support the development of the tool, feedback from service users, managers and key workers was collated in an iterative process.

This particular Outcomes Star was designed for services focused on resolving parental conflict, where the emphasis is on improving effective parenting and child wellbeing. More specifically, it seeks to assess how well parents who are in an intact or separated relationship manage their feelings and any conflicts that arise around child-rearing issues. The tools cover eight areas of parenting life: (1) practical arrangements, (2) routines and stability, (3) money, (4) meeting emotional needs, (5) boundaries and behaviour, (6) relationship skills, (7) managing strong feelings, and (8) wellbeing. These areas (or subscales) are presented to the parents in a star shape, with each subscale measured on a five-stage Journey of Change from 'Stuck' (stage 1) to 'Effective co-parenting' (stages 9–10). The tool is completed collaboratively as part of a conversation between a service user and support practitioner, such as a key worker.

At the time of writing, only one of our key psychometric features – the measure's internal consistency – had been tested. This was analysed in a six-month pilot study shared in confidentiality with EIF by Triangle Consulting. Although the pilot study reported promising internal consistency which would meet our criteria, further analysis is needed both to confirm the internal consistency of the measure and to assess the other psychometric features. While this analysis is planned in the near future, we invite readers to follow the indications proposed by Triangle and not to use the Family Star (Relationships) as the only measure of post-intervention change in impact evaluations. Like other Outcomes Stars, this tool is more likely to be helpful for strengthening the therapeutic relationship and for providing management information on progress. In relation to strengthening the therapeutic relationship, it provides a framework for a conversation between the service user and practitioner, enabling them to reflect on progress together, including discussing the strengths and challenges at hand. In relation to management information, it provides managers with a nuanced picture of needs and progress based on method of completion in that context.

As with other Outcomes Stars, the Family Star (Relationships) is copyrighted under Triangle Consulting Social Enterprise Limited. Organisations intending to use the tool must undergo Star training and obtain licences for their members of staff, all of which are available at a cost.

- For a preview of the tool, see: <https://www.outcomesstar.org.uk/preview-the-stars-resources/>
- For more information, contact Triangle at: <https://www.outcomesstar.org.uk/contact-us/>

5. Guidance on selecting and using outcome measures

To confidently say that what you are delivering is improving the lives of the families involved, you need to conduct an impact evaluation. An important element of this involves selecting one or more validated measures that are consistent with the intervention's theory of change and the population you are working with. Once appropriate measures have been selected, you then need to use these with participants, both before and after the intervention, to assess if the outcomes of parents and children have improved. In this final chapter of the guide, we provide you with recommendations on how best to select and use measures, including practical tips to consider.

5.1 Selecting appropriate measures

First, it is important to note that there is no perfect solution. Often there are trade-offs to consider, for example, between the accuracy and detail of the information gathered by the measure and the time taken to complete it. Indeed, there are a number of factors to work through when deciding what measure is best given the specific context, purpose and resources available for your evaluation. Despite this, we strongly urge you to **always select measures that have been appropriately validated** and are consistent with the theory of change of the intervention you are evaluating (for more information on what we mean by a theory of change, see section 4.3).

Although the psychometric and implementation features that we used to assess the measures included in this guide are not given different weightings, we strongly believe that the most critical criteria are those pertaining to the psychometric features of a measure. Without valid and reliable measures, it is impossible to have confidence in evaluation findings because it is not possible to determine whether any apparent improvement in outcome is due to a true effect of the intervention, or due to an inherent error and bias of the measurement tool used. Hence, while implementation features are important to consider, they are really more secondary.

When selecting measures to use in impact evaluations, we therefore recommend using validated measures and following the guidance set-out below.

1. Select measures based on anticipated outcomes

The measures you select must directly assess the intended outcomes of the intervention or service you are delivering. There is little point in selecting valid and reliable measures if they assess outcomes which are unrelated to your intervention's main goals. For example, if you were evaluating a co-parenting programme and only measured unrelated or distal (longer-term) outcomes, such as general life satisfaction, you would be unable to conclude whether the programme was successful in improving the relationship between co-parents, because this was not something you measured. Similarly, a core element of the programme's theory of change might also be to improve child outcomes through strengthening parents' ability to manage conflict constructively. However, if you only use co-parenting measures as part of your evaluation, then you will not know if the programme has directly benefited children.

Any measures used in an evaluation must therefore be capable of assessing at least one of the anticipated outcomes specified in the intervention's theory of change or logic model.⁴⁴ Pilot studies are particularly useful for determining which outcome measures are most appropriate (and sensitive) to assessing post-intervention change.⁴⁵ In line with this, it may be important to consider measuring outcomes not included in this guide, such as parental mental health. If you have sufficient resources then it is also worth considering using several measures to capture a range of intended outcomes, so long as this can be done without overburdening participants.

Remember that short-term (or possibly medium-term) outcomes are those most likely to be directly impacted by the intervention. Long-term outcomes are important to be aware of but are often more difficult to measure. A well-informed theory of change and logic model can help you to determine what your short- and medium-term outcomes are, as well as how much time is required to see any meaningful changes so that you can plan data collection accordingly.

A practical way of selecting your measure(s) of choice would be to consider what has previously been used in impact evaluations of the programme you are delivering. To gain better insight on how this can be achieved, please see appendix B.

TIP #1: In the context of interventions seeking to improve interparental relationships, remember to select measures that assess both child and parent outcomes, provided the intervention you are delivering intends to impact those outcomes according to its own theory of change. For more details on why it is important to measure changes for children in this context, including links between parental conflict and child outcomes, revisit chapter 3 of this guide.

2. Consider measures that are appropriate for your target participants

The measures you select should be validated for use with a population which resembles the participants receiving your intervention. This is because, even if a measure has been independently validated for a specific population, it may not be valid or reliable (and therefore sensitive to change) when used with a different population. When selecting measures, you should therefore make sure to consider only those that have been designed for and tested with populations that compare well with the participants you are seeking to reach, paying special attention to the following.

- **Age:** A measure which has been validated for use with normally developing teenagers, for example, is likely to be inappropriate (and yield invalid and unreliable results) when used to assess preschool children. Therefore, if you are delivering a programme for parents of 4–8-year-olds which seeks to improve their emotional and behavioural problems, for example, you should not use the MFQ as it has only been validated for 8–18-year-olds. Instead, it would be more appropriate for you to use the SDQ, which is suitable for children of younger ages, including 4–8-year-olds.
- **Level of need:** Measures can be unreliable if they are used for low-need samples when they are designed for and have been tested with higher-need samples, and vice versa. For example, if you are delivering a programme to intact couples in non-abusive conflictive relationships, but have chosen to use a measure of domestic abuse, it is likely that participants will score at the lower limit of this measure. It may therefore be very difficult

⁴⁴ For more information on theories of change and logic models, see Step 1 and Step 2 of the *10 steps for evaluation success* guidance report. See: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

⁴⁵ For more information on pilot studies, please refer to step 5 of the aforementioned guide.

to detect an improvement in couple conflict even if there was one post-intervention because the actual variation in the outcome of interest (that is, conflict) would not be reflected in the scores obtained from the outcome measure used. A domestic abuse measure would have focused on abusive conflict too severe for your target population. In this case, it would have been more appropriate to select a measure enquiring about non-abusive conflict, such as the DAS for intact couples or the OPS for couples with children.

- **Demographics:** An individual's demographic characteristics, including their ethnicity, language and cultural background, will influence many aspects of their life. The measurement of psychological and other outcomes is as a result likely to be influenced by demographic characteristics, such as culture. Measures that have been validated for a particular cultural group, for example, may not necessarily be appropriate for use by another.⁴⁶ Issues with language, wording and the interpretation of certain items may also occur when measures developed for use in one country with a specific group are used in a different setting. A practical way of determining whether the measures you are considering are appropriate for your target population, is to check whether there are any publications describing the use of those measures with that demographic group, paying particular attention to any feedback reported by participants.

TIP #2: Consider the preferred language of your participants when selecting measures.

Many of the most commonly used measures have already been translated and validated in a range of different languages. Therefore, if some of the participants receiving your intervention or service would prefer to complete measures in their native language, it may be very easy for you to arrange this. If, however, measures are unavailable in your participants' preferred language, we would strongly discourage you from translating measures to a different language. Translating validated measures to another language, without testing their psychometric features will not guarantee that the measure remains valid and reliable.

Information on the languages available for the measures included in this guide are presented within the summary reports in appendix A, as well as the full reports available to download from the EIF website.⁴⁷

3. Select an adequate number and appropriate combination of measures that 'fit' together as a whole

Interventions often seek to influence more than one outcome. The Triple P Family Transitions programme for parents experiencing difficulties as a consequence of separation or divorce, for instance, is designed to improve parents' coping skills in managing the transition through separation; increase parental competence and confidence; reduce parental stress; improve co-parenting practices; and strengthen the parent–child relationship.⁴⁸ Ideally an evaluation of this programme would include at least one validated measure for each intended outcome, however, there are risks with using too many measures.⁴⁹ Sometimes interventions also come with their own set of measures, endorsed by the providers themselves. Either way, when deciding what measures to use in your evaluation, you should avoid selecting measures that assess overlapping outcomes, but seek to assess all intended outcomes.

⁴⁶ Mushquash, C.J., and Bova, D.L. (2007). Cross-cultural assessment and measurement issues. *Journal on Developmental Disabilities*, 13(1), 53–65.

⁴⁷ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

⁴⁸ See: <https://guidebook.eif.org.uk/programme/triple-p-family-transitions>

⁴⁹ For more information see: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

Other things you may want to consider when selecting your set of measures include.

- **Brevity:** While there are benefits with using several outcome measures at once, this approach will only work if the overall time taken to complete the entire set of measures is feasible for your participants. This will likely depend on the duration of the intervention being delivered. For example, while it may be acceptable to ask participants taking part in a six-month intervention to complete a one-hour assessment (comprised of three or more outcome measures), the same one-hour assessment would be inappropriate for a week-long intervention due to concerns of overburdening participants. In this instance, provided the measures are consistent with the intervention's theory of change, you may opt for assessing both interparental and child outcomes using the SDQ and DAS-7 (total estimated completion time: 10 minutes), rather than the CBCL/6–18, DAS-32 and CPIC (total estimated completion time: 45 minutes). Respondents who do not complete all the questions within a measure can compromise the quality of the data collected. Therefore, although it is important to avoid overburdening participants, issues of brevity need to be balanced against the validity and reliability of a measure, as longer measures tend to have increased reliability and greater sensitivity to change.
- **Cost:** Many of the most common measures are free to use, but some are copyrighted and come with an associated cost. Before selecting measures, you should therefore familiarise yourself with how much it costs to use them, and consider this in light of the resources available. Of the measures assessed in this guide, only the CBCL/6–18, ECBI and PAM need to be purchased.
- **Required training:** All the measures included in this guide can be completed by either a parent or child; however, there are many other validated measures which need to be administered and/or scored by a researcher or practitioner who has undergone specific training. Arguably this training can increase the precision and sensitivity of the measure, but it also creates an additional burden with respect to the added cost and time required for it. You should therefore consider whether any training is required, in terms of learning how to administer, score or interpret the results of the measures you choose. Of those included in this guide, only the CBCL/6–18 and ECBI need to be interpreted by someone with clinical and/or research expertise, while the SCARED is intended for use by trained clinicians.

TIP #3: Familiarise yourself with the brevity, cost, licensing, ease of scoring and training requirements of each measure. We know that it is difficult to fund robust outcome measurement. The resource required to collect outcome data, due to the training needed to administer or score the measures, as well as the cost of the measures themselves, has been identified as a key barrier to effective practice. Information on these factors has therefore been collected for the measures assessed in this guide, and is included within the full reports available to download from the EIF website.⁵⁰

4. Consider including measures completed by different respondents and collecting administrative or observational data

Many of the measures included in this guide are self-report measures for either children or parents to complete. These measures offer several practical advantages as it is significantly cheaper and less burdensome for individuals to complete measures on their own behalf, compared to alternatives such as observational measures or those reliant on teacher report. Given that they rely on participant perceptions, however, self-report measures can sometimes

⁵⁰ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

lead to inaccurate results, with the quality of data affected by issues such as developmental age, poor literacy or a lack of motivation to provide the true response.⁵¹

For those of you interested in conducting more robust and ambitious studies we would encourage you, where possible, to consider whether your evaluation might benefit from a form of measurement that is independent of the participants taking part in the intervention, which can include the following examples.

- **Measures completed by different respondents** such as parents, teachers and peers, as they are likely to provide different but complementary responses that result in a more detailed assessment of the child in question. For example, while parents may provide access to detailed information about a child's behaviour at home, they will be less capable of doing so in relation to the child's behaviour at school, which is likely better characterised by a teacher or peer. Some of the measures included in this guide, like the SDQ, CBCL/6–18 and ECBI, have versions specifically designed for educators to complete. Due to time and resources constraints, within this guide we have only assessed one version of each selected measure – with the exception of the DAS where we assessed the DAS-7 and DAS-32 (but both are completed by the same respondent). However, within the summary reports (presented in appendix A), we have listed all the available versions of the measure identified from our review of the literature, and specified who the respondent was for each of the versions. We therefore suggest that you consider this information when deciding what measure to select, but please be aware that our assessments are not transferable to other versions of the same measure not assessed in this guide.
- **Administrative data** refers to information collected by government departments and other organisations, for the purposes of registration, transaction and record keeping. In the case of programmes seeking to reduce parental conflict and improve child outcomes, relevant administrative data which you may already have access to includes school absences, detentions and suspensions, educational achievement reports, grade retentions, police callouts, rates of arrest, child protection plans and custody agreements. Given its nature, administrative data does not need to have established psychometric features, although it is important that the data is described in detail and its sources appropriately referenced.

Since there is no perfect measurement, the strengths and limitations of each data source could be balanced through the triangulation of data, enabling you to corroborate the data obtained by using a variety of sources. The use of administrative data alongside validated self-report measures, for example, could increase confidence in the results of your evaluation, so long as you have ensured that this data is consistent with the intervention's theory of change. Including a disproportionate amount of administrative data within your evaluation could also be misinterpreted as 'fishing for outcomes', which is not good practice. Importantly, the additional resource implications regarding the deployment of a varied assessment, also needs to be considered in light of the resources and time available for the evaluation.

5. Conduct further due diligence to finalise your selection of measures

If you have sufficient resources, consider consulting with experts to decide whether the measures you have selected are the most suitable ones. You could, for example, consult with subject-matter experts, specialists in the design of impact evaluations, as well as those with experience using the measures you are considering. It is also worth engaging with practitioners and providers, to gain their perspectives on whether the measures under consideration are practical and appropriately matched to the intervention's intended outcomes and target participants.

⁵¹ Institute for Employment Studies. (2001). *Guide to Measuring Soft Outcomes and Distance Travelled*. <http://webarchive.nationalarchives.gov.uk/20090902143103/http://readingroom.lsc.gov.uk/lsc/SouthEast/distance1.pdf>

Finally, you should endeavour to determine whether the selected measures are viewed as acceptable by those completing them. Responding to personal questions about one's self, relationship or child, can be a daunting and anxiety-producing experience. Moreover, not all measures are suitable for individuals with particular literacy levels or mental health needs. It is therefore your responsibility to determine the suitability of the measures under consideration, paying special attention to the age, capability and personal circumstances of the participants you will be working with. A practical way of determining whether the measures you are considering would be deemed acceptable by your target population, is to ask a representative individual for their opinion and to check whether there are any publications describing problematic use of those measures with particular groups of individuals.

To learn more about the measures you are considering, we would also encourage you to directly contact the developer(s) of the tools themselves.

TIP #4: Consider having a shared set of measures across all interventions delivered in your area. If you are delivering several interventions as part of your service, it would be useful to consider using a shared set of measures. To do this, you should first determine if there are a common set of intended outcomes (such as reduced parental conflict) across the interventions you are delivering, and matching any chosen measures to those outcomes. The benefits of doing this are that results can then be aggregated to allow more robust analysis of the overall effect of the service, as well as enable you to compare the effectiveness of the different interventions you deliver.

Where possible, you should also consider liaising with local areas at both regional and national levels, to encourage their use of the same validated measures. This will allow you to make comparisons with different areas delivering the same or similar interventions, which will help you gain a better understanding of what works best, for whom, where and why.

A note on assessing the impact of online/digital interventions

This practical guide is focused on how to measure interparental relationship and child outcomes in the context of interventions delivered in person. If an intervention is delivered online, there are additional factors to consider in order to select the most appropriate outcome measures.

- Given that digital interventions are highly accessible, participant enrolment is often easier than in the context of interventions delivered in person. This may, however, come with an associated cost regarding the users' motivation and engagement, which has been reported to be lower in interventions delivered online as opposed to face-to-face.⁵² For this reason, it is suggested that the outcome measures selected for the assessment of online interventions are as short as possible. Measures that include less than 10 items can be completed quickly and may minimise the risk of some respondents not completing all items. Despite this, it is important to bear in mind that brevity must always be balanced against the validity and reliability of the measures under consideration.
- Measures assessing the frequency of intervention usage should not be used as outcome measures, as there is mixed evidence to suggest that usage metrics are associated

⁵² Karekla, M., Kasinopoulos, O., Neto, D.D., Ebert, D.D., Van Daele, T., Nordgreen, T., Höfer, S., Oeverland, S., & Jensen, K.L. (2019). Best practices and recommendations for digital interventions to improve engagement and adherence in chronic illness sufferers. *European Psychologist*, 24(1), 49–67.

with an intervention's intended outcomes.⁵³ For instance, while cessation of intervention usage may imply disengagement from the digital intervention, in some cases it might also indicate that users have decided that access to the intervention was no longer needed. Conversely, prolonged usage could signify either positive engagement or excessive dependence on the intervention.

Checklist: considerations for selecting measures

- ☐ The measures have been validated, meaning that they have been tested and shown to be both valid and reliable.
- ☐ Each measure directly assesses at least one intended outcome of the intervention under evaluation. The measures are therefore consistent with the intervention's logic model and theory of change.
- ☐ In the context of interventions seeking to improve interparental relationships, measures assessing both child and interparental/couple outcomes have been considered.
- ☐ For interventions with intended outcomes not included in this guide, such as parental mental health and wellbeing, measures assessing these outcomes have been considered.
- ☐ The measures are appropriate for the target population, meaning that they have been validated with populations that compare well with the participants you are seeking to reach. This has been considered particularly with regards to the participants' age, level of need, demographic characteristics and preferred language.
- ☐ An adequate number and appropriate combination of measures has been selected. This has been decided in light of the time and resources available for the evaluation, while also considering the brevity, cost, licensing, ease of scoring and training requirements of each measure.
- ☐ The acceptability of the measures, as viewed by those completing them, has been considered. Special attention has been paid to the age, capability (such as literacy levels) and personal circumstances (such as mental health needs) of the participants.
- ☐ The full measurement reports, including detailed evidence underpinning all measures included in this guide (and available to download from the EIF website⁵⁴), have been considered. Readers are therefore in a position to make informed decisions that take into consideration the strengths and limitations of each measure. Where necessary, the measurement developer(s) have been contacted.
- ☐ For more robust and ambitious evaluations, a form of measurement that is independent of the participants involved in the intervention has been considered, including measures completed by different respondents (such as teachers and peers) and administrative or observational data.
- ☐ Resources permitting, the selection of measures have been run past a group of experts, including specialists in the design of impact evaluations and those with experience using the measures under consideration.

⁵³ Yardley, L., Spring, B.J., Riper, H., Morrison, L.G., Crane, D.H., Curtis, K., Merchant, G.C., Naughton, F., & Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833–842.

⁵⁴ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

5.2 Using measures appropriately

Once you have selected validated measures that are appropriate for the purpose and context of your evaluation, as suggested in the guidance above, it is important that you use those measures in a suitable manner by taking the following considerations into account.

1. Measures should be used in their entirety

As already mentioned elsewhere in this guide, for measures to be appropriately validated they must undergo a series of stringent statistical tests. Validated measures should therefore be used in their entirety, without changing, adding, deleting or altering the order of any items. While it may be tempting to remove items from a measure in order to reduce its length and shorten its administration time, modifying a validated measure will compromise its ability to detect post-intervention changes accurately and reliably.

In situations where shorter measures are required, it may be feasible to administer only one or a subset of the measure's subscales (for example, the Dyadic Consensus subscale of the DAS-32 or the Problem Behaviour Scale of the CBCL/6–18), so long as this is consistent with the intervention's theory of change and the psychometrics of the subscale have been verified.

TIP #5: When in doubt, always consult the developer(s) of a measure. Generally, it is not recommended that you change the wording of specific items within a measure. There are, however, some exceptions to this rule. When referring to measures assessing interparental relationships, those designed for intact couples many years ago include terms like 'spouse' or 'marriage'. These days, many committed intact couples are not necessarily married but cohabiting; so, changing the term 'spouse' to 'partner' may be more suitable for your sample population. Although this is likely to be permissible, it is important to note that small (seemingly innocent) changes can create bias and affect the quality of the data collected. We therefore recommend that you get in touch directly with the developer(s) to clarify any such queries. You should also report any modifications to a measure explicitly within your evaluation write-up.

2. Measures should be administered in the same way every time and for each group (if applicable)

Bias can be easily introduced during data collection. We therefore recommended that measures are administered in a standardised way, to ensure that data is collected validly and reliably. This means that measures must be administered in the same way every time (for example, face-to-face by a trained practitioner). In cases where there are treatment and control groups (like in randomised controlled trials), we also suggest that these groups are dealt with in a consistent and equivalent manner at all periods of time.

If outcome measurement for the treatment group was different in some way from that of the control group, observed differences between the groups could reflect this differential approach to measurement rather than a true intervention effect, and so results would be biased. Similarly, if outcome data was collected for each group at different periods of time, then the groups would not be truly equivalent at the point of measurement. The estimated programme effects might therefore be confounded by one group being slightly older than the other and possibly exposed to different events, affecting the outcomes measured. Consequently, to ensure consistency in data collection it is important to consider the following.

- **How the outcome measures are presented:** Where possible try to minimise bias by presenting the measures to all respondents in exactly the same way. A practical way of achieving this may be to create a standardised script and encourage respondents to provide honest answers. You may also consider completing an example item with the respondents before they are left alone to complete the measure. In this way, questions or concerns that might affect the measurement can be clarified ahead of time. If you do this, however, make sure to do it for all respondents or it will result in unnecessary discrepancies in measurement.
- **How the outcome data is collected:** Data must be collected in the same way for all participants, particularly with regards to the method, time and duration of measurement.

3. Logistical aspects of research, including ethical approval, informed consent and data collection should be considered before measures are administered

Before using a measure, you must first decide how and when you will use it. For example, if you are interested in assessing the impact of an intervention, it is essential that the selected measures are completed before participants receive the intervention and then again after the intervention is over. To do this you must define what constitutes the beginning and end of the support you are providing, consider when it might be best to collect data, and how you will ensure that the data is collected and stored in a safe and anonymous manner.

According to the General Data Protection Regulation (GDPR),⁵⁵ implemented in May 2018, there are several legal requirements that must be adhered to. Below we provide some examples of things you may need to consider, but **we strongly encourage that you engage fully with the GDPR guidance, as failure to comply may result in severe penalties.**

- **Ethical approval:** Ethical approval is needed for any research project involving human participants, their tissue and/or their data. If you are collecting data from participants for research purposes, you are therefore required to apply for ethical approval from the relevant Research Ethics Committee. This must be sought prior to the start of a research project, as it cannot be granted retrospectively, and conducting research without such approval constitutes a breach of the law. Unless you can demonstrate that your project is not for research purposes, you should apply for ethical approval.⁵⁶
- **Informed consent:** Prior to data collection, you must gain informed written consent from your participants. This is an additional consent to an agreement for participation in the intervention and pertains to the collection, use and sharing of data. Ahead of signing the informed consent, you must ensure that participants understand the purpose of the study, including why this data is being collected and how it will be used. Participants should also be made aware that they are providing their data voluntarily and can withdraw their consent for data use, and ask for the deletion of their data, at any point without their care being compromised.
- **Sensitive personal data:** Anyone collecting, storing and analysing data needs to be aware that data collected from participants, particularly when it is in an identifiable form, including names, dates of birth and home addresses, for example, is considered 'sensitive personal data'. Consequently, when this kind of data is obtained from participants it should be stored in password-protected files (in the case of electronic data) or lockable filing cabinets (in the case of hard copy data). Only anonymised data can be shared with third parties.

⁵⁵ For more information on what GDPR means for health and social care research, see: <https://www.hra.nhs.uk/hra-guidance-general-data-protection-regulation/>

⁵⁶ For more information on applying for ethical approval, including whether this is necessary in your circumstances, see: <https://www.hra.nhs.uk/approvals-amendments/> and <http://www.hra-decisiontools.org.uk/ethics/>

TIP #6: Determine how you will respond if the data collected highlights a particular safeguarding risk. Some of the child measures included in this guide contain sensitive items enquiring about parental physical punishment, self-harm, suicidal ideation and intent. Similarly, a few interparental outcome measures enquire about interpersonal violence or abuse between adults. Where relevant, we have endeavoured to make a note of this within the summary measurement reports in appendix A. In situations of risk or concern, we recommend that those responsible for data collection refer the child or adult onto their GP, a relevant mental health service or a local authority, as appropriate.

Overall, it is important for local areas to be aware of these issues, so they can take appropriate precautions and ensure that data collection methods are properly embedded within local safeguarding procedures. For example, if a child reveals that they are being harmed, abused or neglected, it is your legal responsibility to put in place the safeguarding procedures adhered by your local authority. Moreover, in circumstances where ethical approval has been granted as part of a research project, the type of approval obtained will dictate what your legal responsibilities are. Typically, safeguarding procedures which include procedures for debriefing, signposting and dealing with disclosures need to be resolved as part of the ethical approval process.

When collecting data from participants, it is also important to ensure that you are using the measures in line with any requirements imposed by the developer(s). Some measures, for example, need to be completed face-to-face with the participant, while others must be completed alone by the participant, with limited or no guidance. Wherever possible, we have tried to embed this information within the summary reports in appendix A or the full reports available to download from the EIF website.⁵⁷ For more detailed information, we suggest contacting the developer(s) or searching the measure's website.

Finally, it is also important to introduce outcome measures in a sensitive manner. Most individuals working with families already have the necessary skills to introduce measures in an appropriate and sensitive manner. However, if needed, practitioner training can help to boost confidence and refine skills. Introducing measures sensitively also means familiarising yourself with the selected measures so that you are able to respond to any queries posed by participants. Depending on the type of measures being used, you may also consider testing the measures out with someone other than the programme participant. This can help you understand what it feels like to pose specific questions, and may even help you to better introduce the measure.

4. Use the information collected from your participants to assess impact

Once you have collected the necessary data both before and after the intervention has been delivered, use this data to carry out statistical analyses and assess whether the intervention resulted in any significant changes to the outcomes measured. To do this you may need to draw on help from individuals with analytical experience.⁵⁸ Where possible, and with resources permitting, you should also try to follow-up participants over time, ideally for at least one year after the intervention has concluded. This will enable you to assess whether any identified changes are sustained over time.

⁵⁷ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

⁵⁸ To learn more about relevant statistical tests, see the *10 steps for evaluation success* report, appendix D on statistical power, effect sizes and t-tests, might be of some help. See: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

If you decide to write up your evaluation results, it is important that you acknowledge the authors and developers of the measures used, citing the key papers as appropriate.

TIP #7: Reward participants to try and encourage them to provide follow-up outcome data. Obtaining follow-up outcome data after an intervention has already been delivered is a challenging and time-consuming feat. Nonetheless, acquiring this information can be particularly useful, as it can help to determine whether any positive short-term impacts remain or fade with time. You should therefore make efforts to follow up with as many families as possible. It is common to offer gift vouchers of nominal amounts to thank participants for their involvement in the study. For more information on how to retain study participants, see Step 6 of our *10 steps for evaluation success*⁵⁹ and our report on engaging disadvantaged and vulnerable families.⁶⁰

5. Be aware of the strengths and limitations of the measures you are using

In order to properly score and interpret the results of your outcome measures, it is important to be aware of the strengths and limitations of the measures selected. Many of the most commonly used interparental relationship measures are appropriately validated but not adequate for use with a range of participants. Indeed, many of these measures have been designed and tested using heterosexual married couples and therefore may not be suitable for individuals who are separated or of the same sex. Using an outcome measure with samples for which it has not yet been tested is strongly discouraged. For more details on this, read the second recommendation on selecting appropriate measures (see section 5.1). In addition, we would strongly encourage you to engage with the more detailed evidence underpinning each measure, so that you can better understand the strengths and limitations of the measures selected, and can interpret your results in a more informed manner. The full measurement reports are available to download from the EIF website.⁶¹

TIP #8: Measurement should be independent of any measures used as part of the intervention under evaluation. When measures are used to evaluate the effectiveness of an intervention and also administered as part of it, this limits their ability to accurately assess change as a result of the intervention due to risks of testing effects. Testing effects occur when participants become familiar with the test or testing environment, such that they respond differently when they are next tested. The Parent Daily Report (PDR) presented in appendix A can for example be used to help practitioners identify specific behavioural problems to help tailor parenting advice. When used in this manner, however, the measure becomes part of the actual intervention, meaning that it is no longer sufficiently independent to be used as an assessment tool during rigorous impact evaluations. In these instances, child outcomes should be evaluated using other independent and well-validated instruments.

⁵⁹ See: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

⁶⁰ See: <https://www.eif.org.uk/report/engaging-disadvantaged-and-vulnerable-parents-an-evidence-review>

⁶¹ See: <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

Checklist: considerations for using measures

- ❑ In circumstances where the impact of an intervention is being tested, the selected measures will be completed at least twice, once before participants receive the intervention and then again after the intervention has been delivered.
- ❑ Validated measures will be used in their entirety, without changing, adding, deleting or altering the order of any items.
- ❑ Measures will always be administered in a standardised way, to ensure that data is collected validly and reliably. This will involve considering how the outcome measures are presented to respondents and how the data is collected.
- ❑ In cases where there are treatment and control groups (such as in randomised controlled trials), these groups will be dealt with in a consistent and equivalent manner at all periods of time.
- ❑ Anyone collecting, storing and analysing data will have familiarised themselves with the General Data Protection Regulation (GDPR) and sought ethical approval, where relevant.
- ❑ Local safeguarding procedures are clarified ahead of data collection, so that everyone is aware of how to respond if the data collected highlights a particular safeguarding risk.
- ❑ Measures are used in line with any requirements imposed by the developer(s).
- ❑ Measures are introduced in a sensitive manner. This may involve familiarising yourself with the selected measures and testing it out with someone other than the respondent.
- ❑ In the case of impact evaluations, once the data has been collected, it will be analysed to assess whether the intervention delivered resulted in any significant changes to the outcomes measured.
- ❑ Where possible, resources permitting, participants will continue to be followed up for ideally at least one year after the intervention has been delivered.
- ❑ The strengths and limitations of the measures used are well understood, and allow for an appropriate interpretation of the results.
- ❑ In circumstances where the evaluation results are written up, you will make sure to acknowledge the authors and developers of the measures used, citing the key papers as appropriate.

5.4 Helpful resources

Parental conflict and relationship support

Acquah, D., Sellers, R., Stock, L., & Harold, G. (2017). *Inter-parental conflict and outcomes for children in the contexts of poverty and economic pressure*. London: Early Intervention Foundation. <https://www.eif.org.uk/report/interparental-conflict-and-outcomes-for-children-in-the-contexts-of-poverty-and-economic-pressure>

Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance interparental relationships and improve outcomes for children*. London: Early Intervention Foundation. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>

Lewing, B., Stock, L., & Pote, I. (2017). *Commissioner guide: Reducing parental conflict*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/commissioner-guide-reducing-parental-conflict>

- For more information on EIF's Reducing Parental Conflict Hub, including key 'what works' evidence and tools, see: <https://reducingparentalconflict.eif.org.uk/>
- For more information on the Reducing Parental Conflict Programme, see: <https://reducingparentalconflict.eif.org.uk/about-the-rpc-programme/>
- Details on the eight interventions selected by DWP to be trialled in the UK as part of the national programme can be found at: <https://www.eif.org.uk/files/pdf/cg-rpc-4-3-face-to-face-support-interventions.pdf>

Evaluation

- Asmussen, K., Brims, L., & McBride, T. (2019). *10 steps for evaluation success*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>.
- Martin, J., McBride, T., Brims, L., Doubell, L. Pote, I., & Clarke, A. (2018). *Evaluating early intervention programmes: Six common pitfalls, and how to avoid them*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/evaluating-early-intervention-programmes-six-common-pitfalls-and-how-to-avoid-them>.
- Taylor, S., Drayton, E., & McBride, T. (2019). *Evaluating early help: A guide to evaluation of complex local early help systems*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/evaluating-early-help-a-guide-to-evaluation-of-complex-local-early-help-systems>.

Psychometric features

- Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., & de Vet, H.C.W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63, 737–745.
- Souza, A.C.D., Alexandre, N.M.C., & Ghirardelli, E.D.B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, 26(3), 649–659.
- Terwee, C.B., Mokkink, L.B., Knol, D. L., Ostelo, R.W., Bouter, L.M., and de Vet, H.C. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, 21(4), 651–657.

Selecting and using measures

- Child Outcomes Research Consortium (CORC). *Embedding the use of outcome measures and feedback tools. Guidance for practitioners working with children and young people around their emotional health and mental wellbeing*. <https://www.corc.uk.net/media/1603/embedding-tools-leaflet-final-4web.pdf>
- Child Outcomes Research Consortium (CORC). *Recommendations for using outcome measures*. <https://www.corc.uk.net/media/1487/corc-approach-v3.pdf>
- Law, D., Jones, M., & Wolpert, M. (2015). *Using CYP IAPT feedback and outcome forms to aid clinical practice: Key messages*. Child Outcomes Research Consortium (CORC). https://www.ucl.ac.uk/evidence-based-practice-unit/sites/evidence-based-practice-unit/files/pub_and_resources_resources_for_profs_key_messages.pdf
- Medical Research Council (MRC). Regulatory Support Centre and NHS Health Research Authority. (2018, September). Do I need NHS REC approval?. <http://www.hra-decisiontools.org.uk/ethics/>
- NHS Health Research Authority. (2018, June 13). HRA guidance on the General Data Protection Regulation. <https://www.hra.nhs.uk/hra-guidance-general-data-protection-regulation/>
- NHS Health Research Authority. (n.d.). Approvals and amendments. <https://www.hra.nhs.uk/approvals-amendments/>.
- Pote, I. (2019). *Help us help you: Developing an outcomes framework for addressing parental conflict*. London: Early Intervention Foundation. <https://www.eif.org.uk/blog/help-us-help-you-developing-an-outcomes-framework-for-addressing-parental-conflict>.
- Rhodes, H. (2009). *Knowing what you do works. Measuring your own effectiveness with families, parents and children: a short guide*. London: Family and Parenting Institute. https://tavistockrelationships.ac.uk/images/uploads/download_packs/knowning_what_you_do_works_short_guide_09.pdf
- Wigelsworth, M., Humphrey, N., Black, L., Symes, W., Frearson, K., Ashworth, E., Petersen, K., McCaldin, T., Runacres, J., Demkowicz, O., Pert, K., Tronsco, P., & Anders, L. (2017). *Social, Psychological, Emotional, Concepts of self, and Resilience outcomes: Understanding and Measurement (SPECTRUM): A brief guide to the selection, use, and interpretation of SPECTRUM measures*. London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Evaluation/SPECTRUM/V6_Guidance_document.pdf
- ## Sources of established measures
- For a list of **measures used in evidence-based programmes targeting the interparental relationship**, including measures that specifically examine parental mental health and parenting behaviours (not included in this guide), see: <https://www.eif.org.uk/files/pdf/cg-rpc-3-3-example-measures.pdf>
 - For **measures focused on the early years**, see:
 - EIF's *Foundations for Life* report on what works to support parent–child interactions in the early years includes tables of validated measures, including for children's attachment (p65), children's behaviour (p93), and children's cognitive development (p131). Available at: <https://www.eif.org.uk/report/foundations-for-life-what-works-to-support-parent-child-interaction-in-the-early-years>

- The Education Endowment Foundation (EEF) database of Early Years measures, which includes information on the validity and reliability of measures assessing the language, literacy, numeracy, and social and emotional skills of children aged 0–6 years. Available at: <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/early-years-measure-database/early-years-measures-database>
- A systematic review of parent measures used in impact evaluations of antenatal and early years parenting programmes, including an assessment of the psychometric and implementation properties of these measures. Blower, S.L., Gridley, N., Dunn, A., Bywater, T., Hindson, Z., & Bryant, M. (2019). Psychometric Properties of Parent Outcome Measures Used in RCTs of Antenatal and Early Years Parent Programs: A Systematic Review. *Clinical Child and Family Psychology Review*, 22, 367–387. <https://link.springer.com/article/10.1007/s10567-019-00276-2>
- A systematic review identifying observational measures of parent-child interactions commonly used in impact evaluations of parenting programmes for families with children aged 0-5 years, including an assessment of the psychometric and implementation properties of these measures. Gridley, N., Blower, S., Bywater, T., Whittaker, K., & Bryant, M. (2019). Psychometric Properties of Parent–Child (0–5 years) Interaction Outcome Measures as Used in Randomized Controlled Trials of Parent Programs: A Systematic Review. *Clinical Child and Family Psychology Review*, 22, 253–271. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6478772/>
- For **measures assessing child and adolescent non-academic and essential skills, including mental health and wellbeing outcomes**, see:
 - EEF’s Spectrum Database of measures on non-academic and essential skills with information on their technical quality and ease of implementation. Available at: <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/measuring-essential-skills/spectrum-database/>
 - A review of children’s mental health and wellbeing self-report measures, which have been assessed for the availability of psychometric evidence and suitability for use in routine practice. Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., & Wolpert, M. (2014). Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: a review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health*, 8, 14. <https://capmh.biomedcentral.com/articles/10.1186/1753-2000-8-14>
- For **self-report measures of parental self-efficacy**, see: Wittkowski, A., Garrett, C., Calam, R., & Weisberg, D. (2017). Self-Report Measures of Parental Self-Efficacy: A Systematic Review of the Current Literature. *Journal of Child and Family Studies*, 26(11), 2960-2978. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5646137/>
- Other measurement sources include:
 - The California Evidence-Based Clearinghouse (CEBC) has published a list of reviewed measures assessing child wellbeing, child mental health needs and family attributes, which includes details of how well validated these measures are. See: <https://www.cebc4cw.org/>
 - The Child Outcomes Research Consortium (CORC) includes information on several outcome and experience measures for assessing children and young people’s mental health and wellbeing. See: <https://www.corc.uk.net/outcome-experience-measures/>
 - The Measurement Instrument Database for the Social Sciences (MIDSS) is a repository of over 500 instruments used to collect data from across the social sciences. See: <https://www.midss.org/>
 - The RAND Education Assessment Finder, which provides information about assessments of students’ interpersonal, intrapersonal, and higher-order cognitive competencies. It also includes some useful functionalities to help you filter, review and compare different measures. See: <https://www.rand.org/education-and-labor/projects/assessments/tool.html>

Appendix A: Summary reports of outcome measures

Child outcome measures

- Child Behaviour Checklist (CBCL/6–18)
- Strengths and Difficulties Questionnaire (SDQ)
- Mood and Feelings Questionnaire (MFQ)
- Revised Child Anxiety and Depression Scale (RCADS)
- Screen for Child Anxiety Related Emotional Disorders (SCARED)
- Eyberg Child Behavior Inventory (ECBI)
- Parent Daily Report (PDR)

Interparental outcome measures

- Couples Satisfaction Index (CSI-16)
- Dyadic Adjustment Scale (DAS-32)
- Dyadic Adjustment Scale (DAS-7)
- Golombok Rust Inventory of Marital State (GRIMS)
- Marital Adjustment Test (MAT)
- Relationship Quality Index (RQI)
- Children's Perception of Interparental Conflict Scale (CPIC)
- O'Leary Porter Scale (OPS)
- Parenting Alliance Measure (PAM)
- Parent Problem Checklist (PPC)
- Quality of Co-parental Communication Scale (QCCS)

Child Behaviour Checklist (CBCL/6–18)

113-item parent report of 6–18-year-olds

The Child Behaviour Checklist (CBCL/6–18) is a 113-item parent report measure designed to assess behavioural and emotional problems in children and young people aged 6–18 years. This most recent version of the original measure includes items and subscales aimed at assessing symptoms of anxiety, depression, somatic complaints (that is, physiological symptoms frequently associated with internalising behaviours like anxiety and depression), social problems, thought problems, attention problems, rule-breaking behaviour and aggressive behaviour. The CBCL is part of the Achenbach System of Empirically Based Assessment (ASEBA).

| Psychometric features | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-----------------|-------------------------|----------------|-----------------------|
| | ✓ (Scale) | ✓ (Subscale) | ✓ | ✓ | ✓ |
| Implementation features | Brevity | Availability | Ease of scoring | Used in the UK | |
| | × | × | × | ✓ | |

*Please note that our assessment of this measure is based solely on the English version of the CBCL/6–18. The other versions were not assessed and therefore, it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



Some of the CBCL items contain sensitive content (for example item 18: 'deliberately harms self or attempts suicide'). If an individual raises issues around self-harm, suicide or related issues, they should either be referred to the relevant mental health services or the appropriate safeguarding procedures should be put in place.



**Author(s)/
developer(s)**

Achenbach, T.



**Publication year
for the original
version of the
measure**

1983



**Publication year
for the version
of the measure
assessed**

2001



Type of measure

Parent report
of child

About the measure

| | |
|-------------------------------|--|
| Versions available | <p>There are two other components of the ASEBA – the Teacher's Report Form (TRF) which is completed by teachers for children aged 6–18, and the Youth Self-Report (YSR) for children and young people aged 11–18 years. There is also another version of the CBCL for parents to complete in relation to their 1.5–5-year-olds (the CBCL/1.5-5).</p> <p>In response to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), a version of the CBCL known as the DSM-5-Orientated Scale was also developed and consists of items highly relevant to the diagnostic categories of the DSM-5.</p> |
| Outcome(s) assessed | <p>This measure has been designed to assess internalising and externalising problems in children and adolescents.</p> |
| Subscales | <p>The CBCL/6–18 is comprised of two scales: a problem behaviour scale and a social competence scale.</p> <p>There are eight subscales within the problem behaviour scale, including: anxious/depressed, depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behaviour and aggressive behaviour. These subscales can be grouped into two higher-order factors, known as the internalising and externalising behaviours. By summing up all the problem items, a Total Problems score can also be computed.</p> <p>The CBCL's social competence scale requires parents to report on the child's activities, number of hobbies, involvement with friends, and academic performance.</p> |
| Purpose/primary use | <p>The purpose of this measure is to examine adaptive and maladaptive behaviour, as well as overall functioning in children and young people.</p> |
| Mode of administration | <p>This measure can be completed in person or online.</p> |
| Scoring instructions | <p>This measure has complex scoring instructions involving advanced calculations. The developers suggest that, prior to scoring, each form is checked by a person trained to follow the ASEBA instructions. The CBCL can be hand-scored or computer-scored, with the relevant instructions and scoring materials available at: https://aseba.org/site-and-scoring-licenses/ and https://aseba.org/school-age/.</p> |
| Example item | <p>'Nervous, highstrung, or tense.'</p> |
| Target population | <p>Children aged 6–18 years.</p> |

| | |
|------------------------------------|--|
| Response format | 3-point Likert scale (0 = 'Absent', 1 = 'Occurs Sometimes', 2 = 'Occurs Often'). |
| Language(s) | The CBCL is available in English. According to the developers, it has also been translated into over 100 languages, including Spanish, French and Arabic. Importantly, however, not all forms are available in all languages. For more information, visit: https://aseba.org/translations/ . |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The CBCL is a valid and reliable measure which is sensitive to change in short interventions. • The CBCL assesses both externalising and internalising problems. <p>Limitations:</p> <ul style="list-style-type: none"> • The CBCL is a long and time-consuming measure with 113 items, requiring approximately 15–20 minutes to complete. • There is a cost associated with the use of the CBCL, as well as restrictions over who can purchase it. • The CBCL has complex scoring instructions involving advanced calculations and needs to be scored by someone with clinical and/or research expertise. • According to our review, it does not appear that the CBCL has UK cut-off scores. |
| Link | https://aseba.org/school-age/ |
| Contact details | Sales: mail@aseba.org ; technical support: techsupp@aseba.org |
| Copyright | This measure is copyrighted. For more information, please visit: https://aseba.org/school-age/ |
| Key reference(s) | Achenbach, T., & Rescoral, L.A. (2001). <i>Manual for the ASEBA School-Age Forms & Profiles</i> . Burlington, VT: ASEBA, University of Vermont. |

Strengths and Difficulties Questionnaire (SDQ)

25-item parent/educator report version for 4–17-year-olds

The Strengths and Difficulties Questionnaire (SDQ) is a 25-item measure designed to assess behaviours, emotions and relationships over the last six months in children and young people aged 4–17 years. This original version of the measure is designed for parents or educators to complete, and includes 5 subscales aimed at assessing conduct problems, emotional symptoms, hyperactivity, peer problems and prosocial behaviour.

| Psychometric features | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-----------------|-------------------------|----------|-----------------------|
| | ✓ (Scale) | ✓ (Subscale) | ✓ | ✓ | ✓ |

| Implementation features | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the SDQ for parents or educators of 4–17-year-olds. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- The SDQ is comprised of five subscales, but the internal consistency results for the peer relationship subscale has been found weaker when compared to that of the combined subscales. Consequently, we recommend using the Internalising, Externalising and Total Difficulties scores rather than the individual subscales. For more information, see the 'subscales' section below.
- The developers suggest that the standard SDQ, which asks parents or educators to reflect on a child's behavior in the last six months, should not be used more often than every six months, or there will be overlapping reference periods. To evaluate an intervention shorter than six months, it is therefore best to start with the standard version of the SDQ, and then use a follow-up version, which asks parents to reflect on their child's behaviour in the last month. The follow-up version is available at: [https://www.sdqinfo.com/py/sdqinfo/b3.py?language=Englishqz\(UK\)](https://www.sdqinfo.com/py/sdqinfo/b3.py?language=Englishqz(UK)).



**Author(s)/
developer(s)**

Goodman, R.



**Publication year
for the original
version of the
measure**

1997



Type of measure

Parent report
of child

About the measure

| | |
|-------------------------------|--|
| Versions available | There are four other versions of this measure available, including: a version for parents or educators of 2–4-year-olds, a self-rated version for 11–17-year-olds, another self-rated version for youth aged 18 and over, and an informant report version for youth of that same age. For each of these versions there is also an equivalent follow-up version. For more information, please visit https://www.sdqinfo.com/ . |
| Outcome(s) assessed | This measure has been designed to assess behaviours, emotions and relationships in children and young people. The measure provides both an Internalising and an Externalising score. |
| Subscales | <p>There are five subscales: conduct problems, emotional symptoms, hyperactivity, peer problems and prosocial behaviour.</p> <p>According to the developers, a Total Difficulties score can be generated by summing the scores of all scales except the prosocial scale. The developers also propose summing the scores from the conduct and hyperactivity scales to obtain an Externalising score, and adding the scores of the emotional and peer problems scales to produce an Internalising score.</p> <p>When using a version of the SDQ that includes the 'impact supplement', the items on overall distress and impairment can be summed to generate an Impact (or Impairment) Score.</p> |
| Purpose/primary use | This measure was originally designed to represent strengths and difficulties in children and young people. |
| Mode of administration | This measure can be completed in person or online. |
| Scoring instructions | The SDQ does not need to be scored by someone with specific training or qualifications. This measure can be scored by hand or using the scoring website: https://sdqscore.org/ . |
| Example item | 'Considerate of other people's feelings.' |
| Target population | <p>This measure was originally developed for children aged 4–16 years.</p> <p>However, in June 2014, the developers changed the age range of the standard SDQ from 4–16 to 4–17 years.</p> |
| Response format | 3-point Likert scale (0 = 'Not True', 1 = 'Somewhat True', 2 = 'Certainly True'). |

| | |
|------------------------------------|---|
| Language(s) | The SDQ is available in English. The measure has also been officially translated by the developers into more than 50 languages, with the complete list available here: https://sdqinfo.org/py/sdqinfo/b0.py . |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The SDQ is a valid and reliable measure which is sensitive to change in short interventions. • The SDQ assesses both externalising and internalising problems. • The SDQ is a short (25-item) measure, which is free to access and easy to score (the SDQ is available at: https://www.sdqinfo.com/, with a bespoke scoring website at: https://sdqscore.org/). |
| Link | https://www.sdqinfo.com/ |
| Contact details | youthinmind@gmail.com |
| Copyright | <p>Please note that the SDQ, whether in English or in another language, is a copyrighted document that is not in the public domain. As such, the SDQ may not be modified in any way (for example, by changing the wording of questions, adding questions, or administering only subsets of questions). This is to ensure that the SDQ is fully comparable across studies and settings. Similarly, to ensure high quality and consistency, unauthorised translations are not permitted. Paper versions may be downloaded from the website and subsequently photocopied without charge by individuals or non-profit organisations, provided they are not charging families.</p> <p>Users are not permitted to create or distribute electronic versions for any purpose without prior authorisation from Youth In Mind. If you are interested in making translations or creating electronic versions, you must first contact youthinmind@gmail.com.</p> |
| Key reference(s) | Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. <i>Journal of Child Psychology and Psychiatry</i> , 38(5), 581-586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x |

Mood and Feelings Questionnaire (MFQ)

33-item self-report measure for 8–18-year-olds

The Mood and Feelings Questionnaire (MFQ) is a 33-item self-report measure designed to assess depression in children and young people aged 8–18 years. This original version of the measure includes items indicating how much individuals have felt or acted depressed during the past two weeks.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-------------------------|-----------------|-----------------------|
| Psychometric features | ✓ | ✓ | ✓ | ✓ |
| | Brevity | Availability | Ease of scoring | Used in the UK |
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the 33-item self-report measure. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



Some of the MFQ items contain sensitive content (for instance 'thought about death or dying', 'thought family would be better off without self', 'thought life was not worth living' and 'thought about killing self'). If an individual raises issues around self-harm, suicide or related issues, they should either be referred to the relevant mental health services or the appropriate safeguarding procedures should be put in place.



Author(s)/ developer(s)

Angold, A.,
Costello, E.J.,
Pickles, A., &
Winder, F.



Publication year for the original version of the measure

1987



Publication year for the version of the measure assessed

1995



Type of measure

Child self-report.

About the measure

| | |
|-------------------------------|--|
| Versions available | There are five other versions of this measure available: a 34-item parent report version, a short 13-item child self-report version, an equivalent parent report version, as well as a long and short adult self-report version. |
| Outcome(s) assessed | This measure has been designed to assess depression in children and youth. |
| Subscales | The developers did not provide subscales. However, a recent paper not by the developers, proposed a child-rated suicide-related ideation composite score, which can be computed from the following items: 'thought about death or dying', 'thought family would be better off without self', 'thought life was not worth living' and 'thought about killing self'. |
| Purpose/primary use | This measure was originally developed to assess symptoms of depression in children and adolescents, of clinical and general population samples. |
| Mode of administration | This measure can be completed in person. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found in the developers' letter, available here: https://devepi.duhs.duke.edu/files/2018/03/Dear-potential-MFQ-user.pdf . |
| Example item | 'I felt miserable or unhappy.' |
| Target population | This measure was originally developed for children aged 8–18 years. |
| Response format | 3-point Likert scale (0 = 'Not True', 1 = 'Sometimes True', 2 = 'True'). |
| Language(s) | The MFQ is available in English. According to the official website, the measure has also been translated into the following languages: Arabic-Modern Standard, Arabic-Iraq, Filipino, Finnish, German, Norwegian, Portuguese and Spanish. Further information can be found at: https://devepi.duhs.duke.edu/measures/the-mood-and-feelings-questionnaire-mfq/ . |

| | |
|------------------------------------|---|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The MFQ is a valid and reliable measure which is sensitive to change in short interventions. • It is free to access and easy to score (the measure is available at https://devepi.duhs.duke.edu/measures/the-mood-and-feelings-questionnaire-mfq/, and is scored by summing together the point value responses of each item). <p>Limitations:</p> <ul style="list-style-type: none"> • According to our review, it does not appear that the MFQ has UK cut-off scores. |
| Link | https://devepi.duhs.duke.edu/measures/the-mood-and-feelings-questionnaire-mfq/ |
| Contact details | Brian Small: brian.small@dm.duke.edu |
| Copyright | The developers ask that any published work using the MFQ should cite the authors. |
| Key references | <p>Angold, A., Costello, E.J., Pickles, A. & Winder, F. (1987). The development of a questionnaire for use in epidemiological studies of depression in children and adolescents. London: Medical Research Council Child Psychiatry Unit (unpublished).</p> <p>Angold, A., Costello, E.J., Messer, S.C., Pickles, A., Winder, F., & Silver, D. (1995). The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. <i>International Journal of Methods in Psychiatric Research</i>, 5, 237–249.</p> |

Revised Child Anxiety and Depression Scale (RCADS)

47-item self-report measure for 8–18-year-olds

The Revised Child Anxiety and Depression Scale (RCADS) is a 47-item measure designed to assess symptoms corresponding to anxiety disorders and depression in children and young people aged 8–18 years. The original measure includes six subscales aimed at assessing separation anxiety disorder, social phobia, generalised anxiety disorder, panic disorder, obsessive compulsive disorder and major depressive disorder.

| Psychometric features | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-----------------|-------------------------|----------------|-----------------------|
| | ✓ (Scale) | ✓ (Subscale) | ? | ✓ | ✓ |
| Implementation features | Brevity | Availability | Ease of scoring | Used in the UK | |
| | ✗ | ✓ | ✓ | ✓ | |

*Please note that our assessment of this measure is based solely on the English self-report version of the RCADS, for children and young people aged 8–18 years. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- Some of the RCADS items contain sensitive content (for example item 37: 'I think about death'). If an individual raises issues around self-harm, suicide or related issues, they should either be referred to the relevant mental health services or the appropriate safeguarding procedures should be put in place.
- We found insufficient evidence to establish that the RCADS has good test-retest reliability over short periods of time.
- From our review of the evidence, it appears that the six subscales of the RCADS have a good validity, while that of the total score is questionable. We would therefore encourage you to use the individual subscale scores rather than the total score.



**Author(s)/
developer(s)**
Chorpita, B.F.,
Yim, L., Moffitt, C.,
Umemoto L.A., &
Francis, S.E.



**Publication year
for the original
version of the
measure**

2000



Type of measure

Child self-report.

About the measure

| | |
|---------------------------|--|
| Versions available | There are three additional versions of this measure available, including a parent version (RCADS-P), a shortened 25-item child self-report version and a shortened 25-item parent version. |
|---------------------------|--|

| | |
|----------------------------|---|
| Outcome(s) assessed | This measure has been designed to assess anxiety disorders and depression in children and young people. |
|----------------------------|---|

| | |
|------------------|---|
| Subscales | <p>There are six subscales: separation anxiety disorder (SAD), social phobia (SP), generalised anxiety disorder (GAD), panic disorder (PD), obsessive compulsive disorder (OCD) and major depressive disorder (MDD).</p> <p>The RCADS also yields a Total Anxiety Scale (sum of the five anxiety subscales) and a Total Internalising Scale (sum of all six subscales).</p> |
|------------------|---|

| | |
|----------------------------|--|
| Purpose/primary use | The RCADS measures the reported frequency of various symptoms of anxiety and low mood. It was developed as a revision of the Spence Children's Anxiety Scale (SCAS) in order to correspond to the dimensions of some anxiety disorders reported in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), and also include major depression. In particular, the RCADS was intended to refine the measurement of generalised anxiety disorder (GAD) to reflect core aspects of 'worry'. |
|----------------------------|--|

| | |
|-------------------------------|--|
| Mode of administration | This measure can be completed in person or online. |
|-------------------------------|--|

| | |
|-----------------------------|--|
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found at: https://www.childfirst.ucla.edu/resources/ . |
|-----------------------------|--|

| | |
|---------------------|-------------------------|
| Example item | 'I worry about things.' |
|---------------------|-------------------------|

| | |
|--------------------------|---|
| Target population | This measure was originally developed for children aged 8–18 years. |
|--------------------------|---|

| | |
|------------------------|--|
| Response format | 4-point ordinal scale (0 = 'Never', 1 = 'Sometimes', 2 = 'Often', 3 = 'Always'). |
|------------------------|--|

| | |
|--------------------|--|
| Language(s) | The RCADS is available in English and has also been officially translated into 16 other languages, including French, German, Spanish, Chinese, Dutch, Danish and Greek. The official translations can be found at: https://www.childfirst.ucla.edu/resources/ . |
|--------------------|--|

| | |
|------------------------------------|--|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The RCADS is a valid measure with good internal consistency. • It is free to access and easy to score (the measure is available at: https://www.childfirst.ucla.edu/resources/, with scoring instructions here: https://www.childfirst.ucla.edu/resources/). <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the RCADS has good test-retest reliability over short periods of time. • The RCADS has 47 items and might require more than 15 minutes to be completed. • According to our review, it does not appear that the RCADS has UK cut-off scores. |
| Link | https://www.childfirst.ucla.edu/resources/ |
| Contact details | Bruce Chorpita: chorpita@ucla.edu |
| Copyright | <p>The English and translated versions of the RCADS are copyrighted by Chorpita and Spence. Any use of these instruments implies that the user has read and agreed to the terms of use. Neither the developers nor UCLA are responsible for any third-party use of these instruments by individuals who have not read the RCADS guide or its terms of use. While the RCADS can be used for research purposes, the developers ask, as a professional courtesy, to be informed of this before the study is conducted. Finally, the use of RCADS should always include acknowledgement of the development of the RCADS using appropriate scholarly citations, including the item development contributed by Spence (1997) and extensions by Chorpita et al. (2000).</p> |
| Key reference(s) | <p>Chorpita, B.F., Yim, L.M., Moffitt, C.E., Umemoto L.A., & Francis, S.E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A Revised Child Anxiety and Depression Scale. <i>Behaviour Research and Therapy</i>, 38, 835–855.</p> <p>Spence, S.H. (1997). Structure of anxiety symptoms among children: A confirmatory factor-analytic study. <i>Journal of Abnormal Psychology</i>, 106, 280–297.</p> |

Screen for Child Anxiety Related Emotional Disorders (SCARED)

41-item self-report measure for 8–18-year-olds

The Screen for Child Anxiety Related Emotional Disorders (SCARED) is a 41-item self-report measure designed to assess anxiety disorders in children and young people aged 8–18 years over the past three months. This second version of the original measure includes five subscales aimed at assessing panic/somatic symptoms, generalised anxiety, separation anxiety, social phobia and school phobia.

| Psychometric features | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-----------------|-------------------------|----------|-----------------------|
| | ✓ (Scale) | ✓ (Subscale) | ? | ✓ | ✓ |

| Implementation features | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| | ✓ | ✓ | ✗ | ✓ |

*Please note that our assessment of this measure is based solely on the English self-report version of the SCARED, for children and young people aged 8–18 years. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- The SCARED is comprised of five subscales, but the internal consistency results for the social phobia subscale has been found weaker when compared to that of the combined subscales. Consequently, we recommend using the SCARED total score rather than the individual subscale scores.
- We found insufficient evidence to establish that the SCARED has good test-retest reliability over short periods of time.
- For children aged 8–11 years, the developers recommend that a clinician be present during administration, to explain the questions to the child. Alternatively, it is recommended that the child respond to the questions with an adult present in case there are any queries or concerns.
- Given that the SCARED assesses anxiety disorders over the past three months, we warrant caution when evaluating short interventions lasting less than three months.
- The SCARED is intended for use by trained clinicians.



**Author(s)/
developer(s)**
Birmaher, B.,
Khetarpal, S., Cully,
M., Brent, D., &
McKenzie, S.



**Publication year
for the original
version of the
measure**

1997



**Publication year
for the version
of the measure
assessed**

1999



Type of measure

Child self-report.

About the measure

| | |
|-------------------------------|---|
| Versions available | There are three other versions of this measure available, including the original 38-item version, a shortened 5-item version (to be used in primary care and other community settings), and a 66-item version which has been renamed SCARED-Revised (or SCARED-R). Each of these also have slightly distinct versions for either the parent or child to complete. |
| Outcome(s) assessed | This measure has been designed to assess child anxiety disorders. |
| Subscales | There are five anxiety subscales: panic/somatic symptoms, generalised anxiety, separation anxiety, social phobia and school phobia. |
| Purpose/primary use | This measure was originally designed to be used in clinical and community settings as a screening instrument for child anxiety disorders. |
| Mode of administration | This measure can be completed in person or online. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found at: https://www.pediatricbipolar.pitt.edu/sites/default/files/SCAREDChildVersion_1.19.18.pdf . |
| Example item | ‘When I feel frightened, it is hard to breathe.’ |
| Target population | This measure was originally developed for children aged 8–18 years. |
| Response format | 3-point Likert scale (0 = ‘Not True or Hardly Ever True’, 1 = ‘Somewhat True or Sometimes True’, 2 = ‘Very True or Often True’). |
| Language(s) | The SCARED is available in English but as far as we are aware, the developers did not translate it into other languages. The measure has, however, been translated by people other than the developers into other languages, including into Dutch, French, German, Italian, Portuguese, Spanish, Arabic, Chinese and Thai. |

| | |
|------------------------------------|--|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The SCARED is a valid measure with good internal consistency and it is sensitive to change in short interventions. • It is a free measure (available at: https://www.pediatricbipolar.pitt.edu/sites/default/files/SCAREDChildVersion_1.19.18.pdf). <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the SCARED has a good test-retest reliability over short periods of time. • According to our review, it does not appear that the SCARED has UK cut-off scores. • There are some restrictions over who can purchase the SCARED. • The SCARED is intended for use by trained clinicians. |
| Link | https://www.pediatricbipolar.pitt.edu/resources/instruments |
| Contact details | Boris Birmaher: birmaherb@upmc.edu |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Birmaher, B., Brent, D.A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A replication study. <i>Journal of the American Academy of Child & Adolescent Psychiatry</i> , 38, 1230–1236. |

Eyberg Child Behavior Inventory (ECBI)

36-item parent report of 2–16-year-olds

The Eyberg Child Behavior Inventory (ECBI) is a 36-item measure designed to assess the frequency and severity of disruptive behaviours, as well as the extent to which parents find the behaviours troublesome. The ECBI is specifically focussed on behaviours that take place at home in children and young people aged 2–16 years.

| | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-----------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ (Scale) | ✓ (Subscale) | ? | ✓ | ✓ |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | × | × | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the ECBI. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



We found insufficient evidence to establish that the ECBI has good test-retest reliability over short periods of time.



**Author(s)/
developer(s)**

Eyberg, S.



**Publication year
for the original
version of the
measure**

1978



Type of measure

Parent self-report.

About the measure

| | |
|-------------------------------|--|
| Versions available | Aside from the ECBI, there is an equivalent teacher-report measure, known as the Sutter–Eyberg Student Behaviour Inventory (SESBI; Rayfield et al., 1998), which assess the frequency and severity of disruptive behaviours in school settings. |
| Outcome(s) assessed | This measure has been designed to assess the frequency and severity of externalising behaviours in the home settings, as well as the extent to which parents find these behaviours troublesome. |
| Subscales | This measure is comprised of two subscales: the intensity scale (where the parent indicates how often each behaviour currently occurs) and the problems scale (where the parent indicates whether or not the identified behaviour is a problem). |
| Purpose/primary use | According to our review, the ECBI was originally designed to: <ul style="list-style-type: none">• serve as a brief screening instrument for the differentiation of normal and conduct problems in children and adolescents• provide a sensitive measure of change during the course of treatment• provide a follow-up instrument. |
| Mode of administration | This measure can be completed in person, online or via telephone. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications, but the administration and scoring procedures presented in the ECBI manual should be carefully studied by those administering it. The manual can be purchased here: https://www.parinc.com/Products/Pkey/97 . |
| Example item | ‘Dawdles in getting dressed.’ |
| Target population | This measure can be used for parents of children aged 2–16 years. |
| Response format | A 7-point Likert scale (from 1=‘Never Occurs’ to 7=‘Always Occurs’) and a simple Dichotomous (yes and no) problem scale. |

| | |
|------------------------------------|---|
| Language(s) | The ECBI is available in English and has been translated by the developers into more than 20 languages. Translated versions can be purchased at https://www.parinc.com/Resources/Permissions-and-licensing#99073-eyberg-child-behavior-inventory-ecbi . |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The ECBI is a valid measure with good internal consistency and it is sensitive to change in short interventions. <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the ECBI has good test-retest reliability over short periods of time. • There is a cost associated with the use of the ECBI, and it needs to be interpreted by someone with clinical and/or research expertise. Further details can be found at: https://www.parinc.com/Products/Pkey/97 • According to our review, it does not appear that the ECBI has UK cut-off scores. |
| Link | https://www.parinc.com/Products/Pkey/97 |
| Contact details | https://www.parinc.com/contactus |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Eyberg, S., & Ross, A.W. (1978). Assessment of child behavior problems: The validation of a new inventory. <i>Journal of Clinical Child Psychology</i> , 7, 113–116. https://doi.org/10.1080/15374417809532835 |

Parent Daily Report (PDR)

34-item parent report of 4–10-year-olds

The Parent Daily Report (PDR) is a 34-item measure designed to assess parents' perception of behavioural and emotional problems in children aged 4–10 years. This original version of the measure includes items aimed at assessing negative behaviours displayed by the child within the last 24 hours.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-------------------------|-----------------|-----------------------|
| Psychometric features | ✓ | ✓ | ? | ✓ |
| | Brevity | Availability | Ease of scoring | Used in the UK |
| Implementation features | ✓ | ✓ | ✓ | ? |

*Please note that our assessment of this measure is based solely on the English version of the PDR. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- The PDR has one item examining the use of physical punishment by parents ('parents spank'). If an individual raises issues around parental abuse, the appropriate safeguarding procedures should be put in place.
- We found insufficient evidence to establish that the PDR is a valid measure.
- The PDR can be used to help parents identify specific behavioural problems that can be discussed with practitioners to inform tailored parenting advice. If using the measure in this manner, it becomes part of the intervention, meaning that it is no longer appropriate for use as an assessment tool during rigorous impact evaluations. In these instances, child outcomes can be evaluated using other validated instruments.



Author(s)/ developer(s)

Patterson, G.R.,
Chamberlain, P., &
Reid, J.B.



Publication year for the original version of the measure

1969 (Chamberlain
& Reid, 1987)



Type of measure

Parent report of
child.

About the measure

| | |
|-------------------------------|--|
| Outcome(s) assessed | This measure was originally designed to assess negative behaviours displayed by the child at home but not revealed during observations. |
| Subscales | According to the developers, administration of the PDR yields two scores: a Total Behaviour score, which reflects the number of negative behaviours displayed by the child in the last 24 hours, and a Target Behaviour score, which indicates whether the behaviour was considered stressful or problematic for the parent. |
| Purpose/primary use | This measure was developed as an additional source of information to supplement parent and home observations, and can be used as an outcome measure. The checklist nature of the tool also helps parents identify specific behavioural problems that can be discussed with practitioners and inform tailored parenting advice. |
| Mode of administration | This measure was originally developed to be completed over the phone. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found in the original paper (Chamberlain & Reid, 1987). |
| Example item | 'Noncomplying.' |
| Target population | This measure was originally developed for children aged 4–10 years. From our review of the evidence, it appears that the PDR has been used by the developer with children aged 5–12 years. |
| Response format | Dichotomous scale (simple 'Yes' and 'No' answers). |
| Language(s) | This measure is available in English and as far as we are aware, it has not been translated into other languages. |

| | |
|------------------------------------|---|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The PDR is a reliable measure which is sensitive to change in short interventions. • The PDR is free to access and easy to score (the measure is available at: https://www.oslc.org/pdr/). The Total Behaviour score is the sum of all items, whereas the Target Behaviour score is the sum of all items considered by parents as problematic or stressful. <p>Limitations:</p> <ul style="list-style-type: none"> • From our review, we found insufficient evidence to establish that the PDR is a valid measure. • According to our review, it does not appear that the PDR has been used in the UK and has UK cut-off scores. |
| Link | https://www.oslc.org/pdr/ |
| Contact details | https://www.oslc.org/contact/ |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Chamberlain, P., & Reid, J.B. (1987). Parent observation and report of child symptoms. <i>Behavioral Assessment</i> , 9, 97–109. |

Couples Satisfaction Index (CSI-16)

16-item self-report measure

The Couples Satisfaction Index (CSI-16) is a 16-item measure designed to assess relationship satisfaction of intact (married, cohabiting or dating) couples. This original version of the measure includes items aimed at assessing the presence of problems between individuals and the intensity of such problems.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ | ? | ✓ | ? |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the CSI-16. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- Some versions of the measure show the point values of each response. The authors recommend not showing point values to respondents, but instead using circles to fill in (on pen-and-paper versions) or radio buttons to click (in online surveys).
- We found insufficient evidence to establish that the CSI-16 has good test-retest reliability over short periods of time and is sensitive to change in short interventions.



**Author(s)/
developer(s)**

Funk J.L., & Rogge
R.D.



**Publication year
for the original
version of the
measure**

2007



Type of measure

Self-report.

About the measure

| | |
|-------------------------------|--|
| Versions available | There are two other versions of this measure available: a 32-item version and a 4-item version. |
| Outcome(s) assessed | This measure has been designed to assess relationship satisfaction in intact couples. |
| Subscales | N/A |
| Purpose/primary use | This measure was originally developed to be used with married or cohabiting couples in big studies with large samples and in studies that can only accommodate measures with few items. |
| Mode of administration | This measure can be completed in person or online. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. The full scoring instructions can be found at: https://www.researchgate.net/publication/299432196_The_Couples_Satisfaction_Index_CSI-16 . |
| Example item | 'In general, how often do you think that things between you and your partner are going well?' |
| Target population | This measure was originally developed for married, cohabiting or dating couples. |
| Response format | <p>The CSI-16 is comprised of varying response scales, including ordinal and Likert scales.</p> <ul style="list-style-type: none"> • Item 1 uses a 7-point Likert scale (from 0 = 'Extremely Unhappy' to 6 = 'Perfect'). • Item 2 uses a 6-point ordinal scale (from 0 = 'Never' to 5 = 'All the Time'). • Items 3–6 use a 6-point ordinal scale (from 0 = 'Not at all True' to 5 = 'Completely True'). • Items 7–10 use a 6-point ordinal scale (from 0 = 'Not at All' to 5 = 'Completely'). • Items 11–16 use different 6-point ordinal scales (Item 11: 0 = Boring to 5 = Interesting; Item 12: 0 = Bad to 5 = Good; Item 13: 0 = Empty to 5 = Full; Item 14: 0 = Fragile to 5 = Sturdy; Item 15: 0 = Discouraging to 5 = Hopeful; and Item 16: 0 = Miserable to 5 = Enjoyable). |

| | |
|------------------------------------|---|
| Language(s) | The CSI-16 is available in English but as far as we are aware, the developers did not translate it into other languages. The measure has however been translated by people other than the developers into Persian and Russian. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The CSI-16 is a valid measure with good internal consistency. • The measure is applicable for use with a range of intact couples (such as married, cohabiting, exclusive but not living together, and so on). • It is a short (16-item) measure, which is free to access and easy to score (available at: https://www.researchgate.net/publication/299432196_The_Couples_Satisfaction_Index_CSI-16-16). <p>Limitations:</p> <ul style="list-style-type: none"> • According to our review, we did not find information on the test-retest reliability of the CSI-16. • We also found insufficient evidence to establish that the CSI-16 is sensitive to change in short interventions. |
| Link | N/A |
| Contact details | <p>Email: ronald.rogge@rochester.edu</p> <p>Phone: (585) 273-3270</p> |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Funk, J.L., & Rogge, R.D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. <i>Journal of Family Psychology</i> , 21(4), 572. |

Dyadic Adjustment Scale (DAS-32)

32-item self-report measure

The Dyadic Adjustment Scale (DAS-32) is a 32-item measure designed to assess the relationship quality of intact (married or cohabiting) couples. This original version of the measure includes items and subscales aimed at assessing relationship satisfaction, intimacy, affective expression and the degree to which the couple agrees on matters of importance to the relationship.

| Psychometric features | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-----------------|-------------------------|----------------|-----------------------|
| | ✓ (Scale) | ✓ (Subscale) | ✓ | ✓ | ? |
| Implementation features | Brevity | Availability | Ease of scoring | Used in the UK | |
| | ✓ | ✓ | ✓ | ✓ | |

*Please note that our assessment of this measure is based solely on the English version of the DAS-32. The other versions of this measure were not assessed here and therefore it should not be assumed that they would receive the same rating. The DAS-7 has been assessed separately and is the next measure included in this guide. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- It is unclear whether the DAS-32 is appropriate for use with separated couples. As far as we are aware, the measure was designed for intact couples but has been tested by the developers with both married and divorced couples (Spanier, 1976).
- There is a shorter version of this measure (the DAS-7) with good psychometric values that you might want to consider.



**Author(s)/
developer(s)**

Spanier, G.



**Publication year
for the original
version of the
measure**

1976



Type of measure

Self-report.

About the measure

| | |
|-------------------------------|--|
| Versions available | There are several different versions of the DAS available, including versions with 4, 6, 7, 10 and 14 items (Hunsley et al., 1995, 2001; Sharples & Rogers, 1984). There is also a revised version of the DAS, known as the Revised DAS (RDAS), developed by Busby and colleagues (1995). |
| Outcome(s) assessed | This measure has been designed to assess relationship quality. |
| Subscales | There are four subscales: Dyadic Consensus (the degree to which the couple agrees on matters of importance to the relationship), Dyadic Satisfaction (the degree to which the couple is satisfied with their relationship), Dyadic Cohesion (the degree of closeness and shared activities experienced by the couple) & Affective Expression (the degree of demonstrations of affection and sexual relationships). |
| Purpose/primary use | The DAS-32 was developed for both research and clinical purposes. |
| Mode of administration | This measure can be completed in person or used as a part of an interview. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. The full scoring instructions can be found in the original paper (Spanier, 1976). |
| Example item | 'In general, how often do you think that things between you and your partner are going well?' |
| Target population | This measure was originally developed for couples in intact relationships (cohabiting or married). |

| | |
|------------------------------------|---|
| Response format | <p>The DAS-32 is comprised of varying response scales, including ordinal, Likert and Dichotomous scales.</p> <ul style="list-style-type: none"> • Items 1–15 use a 6-point ordinal scale (from 'Always Agree' to 'Always Disagree') • Items 16–22 use a different 6-point ordinal scale (from 'All the Time' to 'Never') • Item 23 uses a 5-point Likert scale (from 'Every Day' to 'Never') • Item 24 uses a 5-point Likert scale (from 'All of Them' to 'None of Them') • Items 25–28 use a 6-point ordinal scale (from 'Never' to 'More Often'). • Items 29–30 use a Dichotomous (simple 'Yes' or 'No') scale. • Item 31 is rated on a 7-point Likert scale (from 'Extremely Unhappy' to 'Perfect'). • Item 32 asks respondents to choose the most relevant statement from a total of six options. |
| Language(s) | <p>The DAS-32 is available in English. According to our review, the measure has also been translated into several languages including Chinese, Korean, French, Italian and Turkish. However, it is unclear whether these translations have been approved by the developer.</p> |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The DAS-32 is a valid and reliable measure. • The DAS-32 is free to access and easy to score, with scoring instructions available at https://drrebeccajorgensen.com/wp-content/uploads/2011/12/DAS-Scoring-Interpretation.pdf. <p>Limitations:</p> <ul style="list-style-type: none"> • Based on our review, we found insufficient evidence to establish that the DAS-32 is sensitive to change in short interventions. |
| Link | N/A |
| Contact details | N/A |
| Copyright | <p>Based on our review of the evidence, it appears that the developer did not provide information on copyright. The key reference (included below) should be cited when using the measure.</p> |
| Key reference(s) | <p>Spanier, G.B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. <i>Journal of Marriage and the Family</i>, 38, 15–28.</p> |

Dyadic Adjustment Scale (DAS-7)

7-item self-report measure

The Dyadic Adjustment Scale (DAS-7), sometimes referred to as the Abbreviated Dyadic Adjustment Scale (ADAS, Hunsley et al., 2001), is a 7-item measure designed to assess the relationship quality of intact couples. This shortened version of the original DAS-32 (also assessed in this guide) includes items aimed at assessing relationship satisfaction and the degree to which the couple agrees on matters of importance to the relationship.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ | ? | ✓ | ✓ |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the DAS-7. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. The DAS-32 has been assessed separately and is the previous measure in this guide. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- Although the DAS-7 has been used with separated couples (Sharpley & Rogers, 1984), it is unclear whether it is appropriate for use in this context considering that the original DAS (on which this measure was based) was developed for intact couples.
- We did not find information on the test-retest reliability of the DAS-7.



**Author(s)/
developer(s)**

Sharpley C., &
Cross D.



**Publication year
for the original
version of the
measure**

1972



**Publication year
for the version
of the measure
assessed**

1984



Type of measure

Self-report.

About the measure

| | |
|-------------------------------|---|
| Versions available | There are several different versions of the DAS available, including versions with 4, 6, 10, 14 and 32 items. There is also a revised version of the DAS, known as the Revised DAS (RDAS), developed by Busby and colleagues (1995). |
| Outcome(s) assessed | This measure has been designed to assess relationship quality. |
| Subscales | N/A |
| Purpose/primary use | The DAS-7 was developed for both research and clinical purposes. |
| Mode of administration | This measure can be completed in person. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. The total score for the DAS-7 is the sum of the responses to the seven items. |
| Example item | <p>'How often would you say the following events occur between you and your mate?</p> <ul style="list-style-type: none">• Have a stimulating exchange of ideas• Calmly discuss something together• Work together on a project' |
| Target population | This measure was originally developed for intact couples (cohabiting or married), but it has also been used with separated couples (Sharpley & Rogers, 1984). |
| Response format | <p>The DAS-7 is comprised of varying response scales, including both ordinal and Likert scales.</p> <ul style="list-style-type: none">• Items 1–3 use a 6-point ordinal scale (from 5 = 'Always Agree' to 0 = 'Always Disagree').• Items 4–6 also use a 6-point ordinal scale (from 0 = 'Never' to 5 = 'More Often').• Item 7 is rated on a 7-point Likert scale (from 0 = 'Extremely Unhappy' to 6 = 'Perfect'). |
| Language(s) | The DAS-7 is available in English. The measure has also been translated into other languages, including French and Italian, by people other than the developers. |

| | |
|------------------------------------|--|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> Compared to the DAS-32, the DAS-7 is a shorter measure that is available for free, with established validity, internal consistency and sensitivity to change in short interventions. <p>Limitations:</p> <ul style="list-style-type: none"> We did not find information on the test-retest reliability of the DAS-7. If possible, users should assess test-retest reliability in their sample, to improve the evidence-base of the measure. |
| Link | N/A |
| Contact details | N/A |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Sharpley, C. F., & Rogers, H. J. (1984). Preliminary validation of the Abbreviated Spanier Dyadic Adjustment Scale: Some psychometric data regarding a screening test of marital adjustment. <i>Educational and Psychological Measurement</i> , 44, 1045–1049. |

Golombok Rust Inventory of Marital State (GRIMS)

28-item self-report measure

The Golombok Rust Inventory of Marital State (GRIMS) is a 28-item measure designed to assess the overall quality of the relationship between married and cohabiting couples. This original version of the measure includes items aimed at assessing several relationship dimensions, including satisfaction, communication, shared interests, trust and respect.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-------------------------|-----------------|-----------------------|
| Psychometric features | ✓ | ? | ? | ✓ |
| | Brevity | Availability | Ease of scoring | Used in the UK |
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the GRIMS. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



We found insufficient evidence to establish that the GRIMS is a valid measure with good test-retest reliability over short periods of time.



Author(s)/ developer(s)

Rust, J., Bennun, I., Crowe, M., & Golombok, S.



Publication year for the original version of the measure

1986



Type of measure

Self-report.

About the measure

| | |
|-------------------------------|---|
| Outcome(s) assessed | The GRIMS has been designed to assess the quality of the relationship between intact couples through dimensions considered important for a good relationship including communication, shared interests, trust and respect. |
| Subscales | N/A |
| Purpose/primary use | The measure is aimed for use by relationship counsellors and other professionals, to identify the severity of a problem, determine differences in perspectives between partners, and measure relationship change over time. It can also be used in research, to assess the efficacy of different forms of therapy or to investigate the impact of social, psychological, medical or other factors on a relationship. |
| Mode of administration | This measure can be completed in person (with carbonised self-scoring sheets) or online. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions and interpretations can be found within the GRIMS manual: https://www.psychometrics.cam.ac.uk/system/files/documents/GRIMSManual.pdf . |
| Example item | 'My partner is usually sensitive to and aware of my needs.' |
| Target population | The GRIMS was primarily developed to be used with intact (married or cohabiting) heterosexual couples. However, according to the developers, the measure can also be used with couples who are temporarily separated for work or similar reasons, so long as both recognise the other as the primary partner. The developers have also suggested that the measure may be used with homosexual couples, but no standardisation data is currently available for this group. |
| Response format | 4-point ordinal scale (0 = 'Strongly Disagree', 1 = 'Disagree', 2 = 'Agree', 3 = 'Strongly Agree'). |
| Language(s) | According to the developers, the GRIMS is available for online administration in English and has been translated by people other than the developers into a variety of other languages including Malay (Quek et al., 2002). |

| | |
|------------------------------------|---|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The GRIMS has good internal consistency and is sensitive to change in short interventions. • The measure is free to access and easy to score, with scoring instructions found within the GRIMS manual. • The cut-offs for interpretation of the measure are based on a UK sample and therefore standardised to the UK population. • The GRIMS can be used by a range of intact couples, including those that are married or cohabiting. It can also be used with homosexual couples. <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the GRIMS is a valid measure with good test-retest reliability over short periods of time. |
| Link | https://www.psychometrics.cam.ac.uk/services/psychometric-tests/GRIMS |
| Contact details | https://www.psychometrics.cam.ac.uk/contact-us |
| Copyright | <p>According to our review of the evidence, the GRIMS is a copyrighted measure. If you would like to use the GRIMS, please contact the Psychometrics Centre at: https://www.psychometrics.cam.ac.uk/contact-us.</p> |
| Key reference(s) | <p>Rust, J., Bennun, I., Crowe, M., & Golombok, S. (1986). The Golombok Rust Inventory of Marital State, <i>Sexual and Marital Therapy</i>, 1(1), 55–60.</p> |

Marital Adjustment Test (MAT)

15-item self-report measure

The Marital Adjustment Test (MAT), often referred to as the Locke-Wallace Marital Adjustment Scale or the Short Marital Adjustment Test (SMAT), is a 15-item measure designed to assess marital adjustment in married couples. This original version of the measure includes items aimed at assessing the extent of agreement or disagreement between partners on several issues including sex, leisure and money. The measure also explores thoughts and feelings regarding the marriage and one's spouse.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ | ✓ | ? | ✓ |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the MAT. Translated versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- The MAT is one of the earliest relationship quality measures, which is still being used to this day. Questions have however been raised regarding whether some of the language and concepts remain relevant today.
- We found insufficient evidence to establish that the MAT is a valid measure.



**Author(s)/
developer(s)**

Locke, H.J., &
Wallace, K.M.



**Publication year
for the original
version of the
measure**

1959



Type of measure

Self-report.

About the measure

| | |
|-------------------------------|--|
| Outcome(s) assessed | This measure has been designed to assess marital adjustment in married couples. |
| Subscales | N/A |
| Mode of administration | This measure can be completed in person or via telephone interviews. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. The total score is the sum of each point value response, and can range from 0 to 158. |
| Example item | 'State the approximate extent of agreement or disagreement between you and your mate on Handling Family Finances.' |
| Target population | This measure was originally developed for married couples. |
| Response format | <p>The MAT is comprised of varying response scales, including both ordinal and Likert scales.</p> <ul style="list-style-type: none">• Item 1 uses a 7-point Likert scale (from 'Very Unhappy' to 'Perfectly Happy').• Items 2–9 use a 6-point ordinal scale (from 'Always Agree' to 'Always Disagree').• Items 10 asks respondents to select one of three options.• Item 11 uses a 4-point ordinal scale (from 'All of Them' to 'None of Them')• Item 12 asks respondents to select one of two options in relation to themselves and their partner.• Items 13 uses a 4-point ordinal scale (from 'Frequently' to 'Never')• Item 14 and 15 asks respondents to select one of three options and one of four options, respectively. |
| Language(s) | The MAT is available in English but as far as we are aware, the developers did not translate the MAT into other languages. The measure has, however, been translated into Persian by people other than the developers. |

| | |
|------------------------------------|--|
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The MAT is a reliable measure which is sensitive to change in short interventions. • The MAT is free and easy to score, with the total score being the sum of the individual responses. <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the MAT is a valid measure. • The MAT is one of the earliest relationship quality measures (developed in the 1950s), with some items no longer appropriate to modern relationships (Freeston & Plechaty, 1997; Graham et al., 2011). • According to some authors (Shapiro & Gottman, 2005), the MAT is inappropriate for assessing outcomes in most marital interventions. The reason for this contention is that there are two ways of getting a high score on this scale: <ul style="list-style-type: none"> (1) to rate one's relationship as closer to perfectly happy (the scale ranges from 0 [very unhappy] to 15 [happy] to 35 [perfectly happy]); (2) to be conflict averse and not disagree very much, rating the following items closer to agree (on a scale that ranges from always disagree to always agree): handling family finances, matters of recreation, demonstrations of affection, friends, sex relations, conventionality, philosophy of life, and ways of dealing with in-laws. |
| Link | N/A |
| Contact details | N/A |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Locke, H.J., & Wallace, K.M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. <i>Marriage and Family Living</i> , 21, 251–255. |

Relationship Quality Index (RQI)

6-item self-report measure

The Relationship Quality Index (RQI) is a 6-item measure designed to assess the quality of relationships in married and cohabiting couples. This version of the measure is an adaptation of the original Quality of Marriage Index (QMI) designed only for married couples, and includes items aimed at assessing the presence of problems between individuals and the intensity of such problems.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-------------------------|----------------------|-------------------------|-----------------|-----------------------|
| Psychometric features | ✓ | ? | ✓ | ✓ |
| | Brevity | Availability | Ease of scoring | Used in the UK |
| Implementation features | ✓ | ✓ | ✗ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the RQI. The other versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



Based on our review of the evidence, we did not find information on the test-retest reliability of the RQI.



**Author(s)/
developer(s)**

Norton, R.



**Publication year
for the original
version of the
measure**

1983



Type of measure

Self-report.

About the measure

| | |
|------------------------------------|--|
| Outcome(s) assessed | This measure has been designed to assess relationship quality. |
| Subscales | N/A |
| Purpose/primary use | This measure was originally designed to be used by clinical researchers who are interested in screening for relationship satisfaction. |
| Mode of administration | This measure can be completed in person or over the phone. |
| Scoring instructions | This measure has complex scoring instructions involving advanced calculations, however, it does not need to be scored by someone with specific training or qualifications. More information can be found in the original paper (Norton, 1983). |
| Example item | 'We have a good relationship.' |
| Target population | This measure can be used with married and cohabiting couples. |
| Response format | The RQI is comprised of a 7-point Likert scale (from 1 = 'Strongly Disagree' to 7 'Strongly Agree') and a 10-point ordinal scale (from 1 = 'Extremely Low' to 10 = 'Extremely High'). |
| Language(s) | This measure is available in English but as far as we are aware, the developers did not translate the RQI into other languages. The measure has, however, been translated into German by people other than the developer. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The RQI is a valid measure with good internal consistency and it is sensitive to change in short interventions. • It is a short (6-item) measure that is free to access, although permission is required from the publisher to use it. <p>Limitations:</p> <ul style="list-style-type: none"> • From our review, we did not find information on the test-retest reliability of the RQI. • The RQI has complex scoring instructions involving advanced calculations. |
| Link | N/A |

| | |
|-------------------------|--|
| Contact details | N/A |
| Copyright | Based on our review of the evidence, it appears that the RQI is protected by copyright held by John Wiley and Sons, Ltd. Permission to use or reproduce the RQI should be forwarded to the publisher at permissionsUK@wiley.com . In addition, the key reference (included below) should be cited when using the measure. |
| Key reference(s) | Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. <i>Journal of Marriage and the Family</i> , 45(1), 141–151. |

Children's Perception of Interparental Conflict Scale (CPIC)

48-item self-report measure for 9–17-year-olds

The Children's Perception of Interparental Conflict Scale (CPIC) is a 48-item self-report measure designed for children and young people aged 9–17 years, to assess their views of parental conflict and child adjustment. The original version of the CPIC assessed here includes subscales aimed at assessing frequency, intensity, resolution, content, perceived threat, coping efficacy, self-blame, triangulation and stability.

| | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ | ✓ | ✓ | ? | ? |
| | (Scale) | (Subscale) | | | |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | × | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the CPIC. Translated versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- Some of the CPIC items contain sensitive content (for example 'my parents have pushed or shoved each other during an argument' and 'when my parents argue I worry that one of them will get hurt'). If an individual raises issues around interparental violence, the appropriate safeguarding procedures should be put in place.
- According to the developers, the CPIC can be used with children from separated or divorced families, but in these cases the wording of the questions (that is, current or past conflict) should be made explicit.
- We found insufficient evidence to establish that the CPIC is a valid measure which is sensitive to change in short interventions.



Author(s)/ developer(s)

Grych, J.H., Seid,
M., & Fincham, F.D.



Publication year for the original version of the measure

1992



Type of measure

Child self-report.

About the measure

| | |
|-------------------------------|---|
| Outcome(s) assessed | This measure has been designed to assess children's views of multiple dimensions of parental conflict and child adjustment. |
| Subscales | <p>Originally, the measure was comprised of 9 subscales: frequency, intensity, resolution, content, perceived threat, coping efficacy, self-blame, triangulation and stability.</p> <p>The developers derived three broad-factor scales using Exploratory Factor Analysis and Confirmatory Factor Analysis:</p> <ol style="list-style-type: none"> 1. Conflict Properties reflects how often conflicts occur and the level of hostility and resolution. It is comprised of the frequency, intensity and resolution subscales. 2. Threat indicates the degree to which children feel threatened and able to cope when marital conflict occurs. It is comprised of the threat and coping efficacy subscales. 3. Self-Blame assesses the frequency of child-related conflict and the degree to which children blame themselves for marital conflict. It is comprised of the content and self-blame subscales. <p>The stability and triangulation subscales can be used as independent subscales.</p> |
| Purpose/primary use | This measure was originally designed to assess particular dimensions of marital conflict that might lead to child adjustment problems, and to obtain children's perspective on the degree of conflict to which they are exposed. The developers' motivation for developing this measure was based on the fact that parent-report measures often underestimate children's exposure to conflict. |
| Mode of administration | This measure can be completed in person or online. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found in the original paper (Grych et al., 1992). |
| Example item | 'I never see my parents arguing or disagreeing.' |
| Target population | This measure was originally developed for children aged 9–17 years. Further studies have shown that the measure is also appropriate for assessing specific aspects of parental conflict in late adolescents and emerging adults, aged 18–25 years. |

| | |
|------------------------------------|--|
| Response format | 3-point Likert scale (1 = 'True', 2 = 'Sort of True', 3 = 'False'). |
| Language(s) | The CPIC is available in English but as far as we know, the developers did not translate this measure into other languages. The measure has however been translated into German, Greek, Italian, Portuguese, Spanish and Turkish by people other than the developers. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The CPIC is a reliable measure. • It assesses the child's perception of parental conflict; an important outcome that is rarely assessed by other measures. • The CPIC is freely available and easy to score, with scoring instructions found in the paper by Grych et al. (1992). <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the CPIC is a valid measure which is sensitive to change in short interventions. • The CPIC has 48 items and may require more than 15 minutes to complete. |
| Link | N/A |
| Contact details | John Grych: john.grych@marquette.edu |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Grych, J.H., Seid, M., & Fincham, F.D. (1992). Assessing marital conflict from the child's perspective: The Children's Perception of Interparental Conflict Scale. <i>Child Development</i> , 63, 558–572. |

O'Leary Porter Scale (OPS)

10-item self-report measure

The O'Leary Porter Scale (OPS) is a 10-item measure designed to assess overt hostility in intact couples. This original version of the measure includes items aimed at assessing the frequency of overt hostility (such as quarrels, sarcasm, physical abuse) observed by the child.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ | ? | ? | ✓ |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the OPS. Translated versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- The OPS includes two items assessing the frequency of physical and verbal hostility displayed by parents in front of the child. If an individual raises issues around interparental violence, the appropriate safeguarding procedures should be put in place.
- We found insufficient evidence to establish that the OPS is a valid measure with good test-retest reliability over short periods of time.



**Author(s)/
developer(s)**

O'Leary K.D., &
Porter B.



**Publication year
for the original
version of the
measure**

1980



Type of measure

Self-report.

BACK TO LIST OF ALL MEASURES

About the measure

| | |
|------------------------------------|---|
| Outcome(s) assessed | This measure has been designed to assess the frequency of overt hostility witnessed by a child. |
| Subscales | N/A |
| Mode of administration | This measure can be completed in person. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. The first nine questions are scored positively (never = 4; very often = 0) while the final question is scored negatively (never = 0; very often = 4). Once all questions have been scored, the individual scores are summed together. This total overt hostility score therefore ranges from 0–40, with lower scores indicating a greater frequency of overt hostility witnessed by the child. |
| Example item | 'In every normal marriage there are arguments. What percentage of the arguments between you and your spouse would you say take place in front of this child?' |
| Target population | This measure was originally developed for married couples with children; it is now also used with cohabiting and divorces couples. |
| Response format | A 6-point ordinal scale (from 1 = 'Never' to 6 = 'Very Often'). |
| Language(s) | The OPS is available in English but as far as we are aware, the developers did not translate the OPS into other languages. The measure has, however, been translated by people other than the developers into other languages including Spanish. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The OPS has good internal consistency and is sensitive to change in short interventions. • It is a short (10-item) measure, which is free to access and easy to score. <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the OPS is a valid measure with good test-retest reliability over short periods of time. |
| Link | N/A |

| | |
|-------------------------|--|
| Contact details | Daniel O'Leary: Daniel.Oleary@stonybrook.edu |
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Porter, B., & O'Leary, K.D. (1980). Marital discord and childhood behavior problems. <i>Journal of Abnormal Child Psychology</i> , 8(3), 287–295. |

Parenting Alliance Measure (PAM)

20-item self-report measure

The Parenting Alliance Measure (PAM), originally called the Parenting Alliance Inventory (PAI), is a 20-item measure designed to assess the strength of the parenting alliance. This original version of the measure is for parents of children aged 1–19 years, and includes items aimed at assessing how cooperative, communicative and mutually respectful parents are when caring for their children.

| | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-----------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ (Scale) | ✓ (Subscale) | ? | ? | ✓ |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✗ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the PAM. Translated versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



We found insufficient evidence to establish that the PAM is a valid measure with good test-retest reliability over short periods of time.



Author(s)/ developer(s)

Abidin, R.R., &
Brunner, J.F.



Publication year for the original version of the measure

1995



Type of measure

Self-report.

About the measure

| | |
|------------------------------------|--|
| Outcome(s) assessed | This measure has been designed to assess the perceived strength of the parenting alliance. |
| Subscales | There are two subscales: Respect; Communication and Teamwork. |
| Purpose/primary use | According to the official website, the PAM can be used as a screening and diagnostic instrument for family counselling, evaluating joint custody, identifying dysfunctional parenting skills, and assessing the impact of interventions. |
| Mode of administration | This measure can be completed in person. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found in the original paper (Abidin & Brunner, 1995). |
| Example item | 'When there is a problem with our child, we work out a good solution together.' |
| Target population | This measure was originally developed for intact couples (married or cohabiting) with children (1–19 years old). According to the official website, the measures is appropriate for a variety of other co-parents, including those that are separated or divorced. |
| Response format | 5-point Likert scale (from 1 = 'Strongly Disagree' to 5 = 'Strongly Agree'). |
| Language(s) | The PAM is available in English but as far as we are aware, the developers did not translate the PAM into other languages. The measure has, however, been translated into Italian by people other than the developers. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The PAM has good internal consistency and is sensitive to change in short interventions. • The PAM is a short (20-item) measure that is easy to score. <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the PAM is a valid measure with good test-retest reliability over short periods of time. • There is a cost associated with the use of the PAM. |

| | |
|-------------------------|---|
| Link | https://www.parinc.com/Products/Pkey/304 |
| Contact details | N/A |
| Copyright | The measure is copyrighted and can be purchased at: https://www.parinc.com/Products/Pkey/304 . |
| Key reference(s) | Abidin, R.R., & Brunner, J.F. (1995). Development of a parenting alliance inventory. <i>Journal of Clinical Child Psychology</i> , 24(1), 31–40. |

Parent Problem Checklist (PPC)

16-item self-report measure

The Parent Problem Checklist (PPC) is a 16-item measure designed to assess parental conflict over child-rearing issues over the past four weeks. This original version of the measure includes subscales aimed at assessing the presence and intensity of problems between co-parents, regardless of whether they are together or apart.

| | Internal consistency | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ | ? | ? | ✓ |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✓ | ✓ | ✓ |

*Please note that our assessment of this measure is based solely on the English version of the PPC. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



- We found insufficient evidence to establish that the PPC is a valid measure with good test-retest reliability over short periods of time.
- From our review of the evidence, we found that the psychometric features were assessed only using Australian samples. Due to the differences between Australian and UK contexts, we warrant caution when interpreting the evidence.



**Author(s)/
developer(s)**

Dadds, M.R., &
Powell, M.B.



**Publication year
for the original
version of the
measure**

1991



Type of measure

Self-report.

About the measure

| | |
|------------------------------------|--|
| Outcome(s) assessed | This measure has been designed to assess parental conflict over child-rearing issues. It assesses parents' ability to cooperate when performing parenting duties, and was devised to examine the effect of co-parenting conflict on child adjustment problems. |
| Subscales | There are two subscales: Problem Scale and Extent Scale. According to the developers, the Problem Scale reflects the number of disagreements between parents, while the Extent Scale reflects the intensity of the conflicts. |
| Mode of administration | This measure can be completed in person. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found in the original paper (Dadds & Powell, 1991). |
| Example item | 'Disagreement over household rules (such as bedtime, play areas).' |
| Target population | This measure was originally developed for parents of children aged 0–18 years. |
| Response format | 7-point Likert scale (from 1 = 'Not at All', to 7 = 'Very Much'). |
| Language(s) | The PPC is available in English but we are not aware of any translated versions of this measure. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none">• The PPC has good internal consistency and is sensitive to change in short interventions.• The PPC is a short (16-item) measure that is free to access and easy to score. <p>Limitations:</p> <ul style="list-style-type: none">• We found insufficient evidence to establish that the PPC is a valid measure with good test-retest reliability over short periods of time.• Based on our search of the evidence, the PPC has primarily been used in the Australian context, and there are no available cut-off scores for the UK population. |
| Link | N/A |
| Contact details | Mark R. Dadds: mark.dadds@sydney.edu.au |

| | |
|-------------------------|--|
| Copyright | Based on our review of the evidence, it appears that the developers did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Dadds, M.R., & Powell, M.B. (1991). The relationship of interparental conflict and global marital adjustment to aggression, anxiety, and immaturity in aggressive and nonclinic children. <i>Journal of Abnormal Child Psychology</i> , 19, 553–567. |

Quality of Co-parental Communication Scale (QCCS)

10-item self-report measure

The Quality of Co-parental Communication Scale (QCCS), sometimes referred to as the Discuss and Share Decision-Making Scale, is a 10-item measure designed to assess co-parental communication in separated or divorced couples. This original version of the measure includes two subscales aimed at assessing conflict and support.

| | Internal consistency | | Test-retest reliability | Validity | Sensitivity to change |
|-----------------------|----------------------|-----------------|-------------------------|----------|-----------------------|
| Psychometric features | ✓ (Scale) | ✓ (Subscale) | ? | ✓ | ? |

| | Brevity | Availability | Ease of scoring | Used in the UK |
|-------------------------|---------|--------------|-----------------|----------------|
| Implementation features | ✓ | ✓ | ✓ | ? |

*Please note that our assessment of this measure is based solely on the English version of the QCCS. Translated versions of this measure were not assessed and therefore it should not be assumed that they would receive the same rating. For a more thorough understanding of this measure, including a full list of the references used to assess this measure against the predefined criteria, please refer to <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>.



We found insufficient evidence to establish that the QCCS has good test-retest reliability over short periods of time and is sensitive to change in short interventions.



**Author(s)/
developer(s)**

Ahrons, C.



**Publication year
for the original
version of the
measure**

1981



Type of measure

Self-report.

About the measure

| | |
|------------------------------------|--|
| Outcome(s) assessed | This measure has been designed to assess co-parental communication in separated and divorced couples with children. |
| Subscales | The measure has two subscales: Conflict and Support |
| Purpose/primary use | This measure was originally developed within a study aimed at assessing the relationship between parents one year following divorce. |
| Mode of administration | This measure can be completed in person. |
| Scoring instructions | This measure can be scored by hand and does not need to be scored by someone with specific training or qualifications. Scoring instructions can be found in the original paper (Ahrons, 1981). |
| Example item | 'When you and your former spouse discuss parenting issues, how often does an argument result?' |
| Target population | This measure was originally developed for separated and divorced couples. |
| Response format | A 5-point Likert scale (from 1 = 'Never' to 5 = 'Always'). |
| Language(s) | The QCCS is available in English but as far as we are aware, the developers did not translate the QCCS into other languages. The measure has, however, been translated into Japanese by people other than the developers. |
| Strengths & limitations | <p>Strengths:</p> <ul style="list-style-type: none"> • The QCCS is a valid measure with good internal consistency. • It is a short (10-item) measure, which is freely available and easy to score. <p>Limitations:</p> <ul style="list-style-type: none"> • We found insufficient evidence to establish that the QCCS is sensitive to change in short interventions and has good test-retest reliability over short periods of time. • According to our review, there is limited evidence that the QCCS has been substantially used in the UK. |
| Link | N/A |

| | |
|-------------------------|---|
| Contact details | Dan Lainer-Vos: lainer-vos@usc.edu |
| Copyright | Based on our review of the evidence, it appears that the developer did not provide information on copyright. The key reference (included below) should be cited when using the measure. |
| Key reference(s) | Ahrons, C. R. (1981). The continuing coparental relationship between divorced spouses. <i>American Journal of Orthopsychiatry</i> , 51, 415–428. |

Appendix B: Guiding you to identify measures used in impact evaluations

A practical way of selecting your measure(s) of choice is to consider what measures have previously been used in impact evaluations of the programme(s) you are delivering. To gain better insight on how this can be achieved, we provide a worked example applied to the 'Parents Plus Parenting when Separated' intervention.

Parents Plus Parenting when Separated

The Parents Plus Parenting when Separated is one of the eight face-to-face interventions being trialled as part of the national Reducing Parental Conflict Programme. It is a targeted-selective intervention specifically designed to address the needs of separated custodial and non-custodial parents in an Irish context. The intervention is delivered over six-weekly group sessions, each lasting two hours. The sessions cover a range of topics, including: co-parenting, helping your children cope, being a live-away or resident parent, conflict management, and coping in the long term.

To identify measures used in impact evaluations of a given programme, you can take the following steps:

1. Search for relevant studies using the EIF Guidebook¹
2. Search through the relevant programme's website, if one exists
3. Use Google Scholar as a search engine
4. Consider impact evaluations identified by other clearinghouses
5. Extract information from the impact evaluations gathered in steps 1–4.

In some cases, not all steps will yield a result (for instance, not all programmes are available on the EIF Guidebook, nor do all programmes have a bespoke website). It is therefore important that you follow all of these steps, which have been described in more detail below.

Step 1: Searching the EIF Guidebook

The EIF Guidebook provides information about early intervention programmes that have been evaluated and shown to improve outcomes for children and young people. It also includes an 'Improving interparental relationships' spotlight set,² where you can search for relevant programmes. If your programme of interest is listed on the EIF Guidebook, you can then use the following strategy to gather information on its impact evaluations.

¹ See: <https://guidebook.eif.org.uk/>

² See: <https://guidebook.eif.org.uk/search?sets%5B%5D=%25%22improving-interparental-relationships%22%25>

Finding relevant impact evaluations

Once you have identified the relevant programme, open the programme page and use the menu on the left-hand side to select the 'About the evidence' page. Here, you will find details of the impact evaluations reviewed by EIF for this particular programme.

In the case of the Parents Plus Parenting when Separated intervention, for example, the programme is in the EIF Guidebook. For details on the evidence underpinning the programme, select 'About the evidence'.

FIGURE B.1

EIF Guidebook entry providing information about Parents Plus Parenting when Separated



Identifying the measures used in impact evaluations

Within the EIF Guidebook, on the 'About the evidence' page or the selected programme, you will find a summary of all the impact evaluations (or studies) assessed by EIF. To find out what measures have been used in these impact evaluations, click on 'About study 1' and scroll to the 'Measures' section. In this section you will find a list of all the outcome measures used in that study. If there is more than one study linked to the programme you are interested in, don't forget to apply this same process to the other studies.

In the case of Parents Plus Parenting when Separated, for instance, the EIF Guidebook only includes one study of this programme – the Keating et al. (2015)³ study – or study 1, as shown in figure B.2.

When selecting the 'More about study 1' icon, you will then be able to find that the study used four measures: the Kansas Parental Satisfaction Scale, the Strengths and Difficulties Questionnaire, the Quality of Co-parental Communications Scale, and the Mental Health Inventory.

³ Keating, A., Sharpy, J., Murphy, M., Rooney, B., Carr, A. (2016). An evaluation of the Parents Plus – Parenting When Separated programme. *Clinical Child Psychology and Psychiatry*, 21(2), 240–254.

FIGURE B.2

EIF Guidebook entry providing information about Study 1 of the Parents Plus Parenting when Separated

The screenshot shows the EIF Guidebook interface. At the top, there's a purple header with the EIF logo and navigation links. Below the header, a breadcrumb trail reads 'GUIDEBOOK > Parents Plus Parenting when Separated'. A text box states: 'This programme is underpinned by one study with a level 2+ rating, hence the programme receives a level 2+ rating overall.' Below this, a section titled 'Study 1' contains a table with the following details:

| | |
|-----------------|--|
| Citation: | Keating et al., 2015 |
| Design: | Cluster RCT |
| Country: | Ireland |
| Sample: | 16 sites (Child and Family Mental Health Services, and Child and Family voluntary care agencies), including 161 separated families, with children 3 years or older |
| Timing: | Pre-test; post-test (six weeks later) |
| Child outcomes: | Improved child behaviour |
| Other outcomes: | Improved parenting satisfaction Improved parental adjustment Reduced interparental conflict |
| Study rating: | 2+ |

Below the table is a button labeled 'More about study 1' with a downward arrow.

Accessing impact evaluations

If you are interested in reading the actual study that assesses the impact of the given programme, the link to the online publication is immediately presented once you click the 'More about study 1' icon. Sometimes, published studies can only be accessed with an academic account; however, many papers are free to download on Google Scholar⁴ and ResearchGate.⁵

Step 2: Searching the programme website

Identifying the programme website

If the programme you are interested in is in the EIF Guidebook, you can find the website's official link (if it exists) at the bottom of the 'About the programme' webpage, under 'Contact details'. If the programme is not in the Guidebook, you should try and find the programme's official website by searching for it via Google or an alternative search engine.

Because Parents Plus Parenting when Separated has an official website and is in the EIF Guidebook, the link to their website is reported at the bottom of the 'About the programme' webpage.⁶ The same website was also among the first results when we searched for the programme on Google.⁷

⁴ See: <https://scholar.google.com/>

⁵ See: <https://www.researchgate.net/login>

⁶ See: <https://guidebook.eif.org.uk/programme/parents-plus-parent-when-separated#about-the-programme>

⁷ See: <https://www.parentsplus.ie/programmes-about/parenting-when-separated/>

Finding relevant impact evaluations

Once on the programme's website, navigate to the pages containing the relevant impact evaluations. In general, impact evaluations will be listed in the 'Research', 'Evidence' or 'Studies' section of a website. For instance, the website of Parents Plus Parenting when Separated listed one impact evaluation under the 'Research' section. The study they refer to is the same as that which we included in the Guidebook (the Keating et al. (2015) study).

Accessing impact evaluations

If the programme website provides only references to the relevant impact evaluations, instead of the link to the publication, you can search the titles of the studies on Google Scholar or ResearchGate. Before downloading the publications, make sure that the studies are different from those identified in step 1 of this guide (using the EIF Guidebook).

In the case of Parents Plus Parenting when Separated, for example, the programme website does provide a link to the pdf file of the relevant study.⁸

Step 3: Using Google Scholar

Finding relevant impact evaluations

The next step is to use the advanced search option of Google Scholar. To do this, open the Google Scholar link and select the icon with the three lines in the top left-hand corner. Then select 'Advanced Search'.

Next, you will need to:

- type the programme name into the 'with exact phrase' box
- type the following string of terms into the 'with at least one of the words' box: impact evaluate evaluation intervention result effect effective efficacy efficacious trial study (note: these terms will help to narrow your search, so that you are more likely to identify impact evaluations of the relevant programme).

To find the impact evaluation of Parents Plus Parenting when Separated, for instance, we completed the search form as shown below.

Depending on how much time you have, you might want to check the first five pages of results in order to identify relevant entries. If you find an impact evaluation that looks relevant, you can read the 'Abstract' and the 'Methodology' or the 'Measures' sections of the article to decide whether the study contains relevant information.

For instance, when we used the Google Scholar advanced search engine, we found 31 results. However, when we went through them individually, we found that only one of these search results was an impact evaluation of the relevant programme (the same study identified in steps 1 and 2 above, the Keating et al. (2015) study).

⁸ See: <https://www.parentsuplus.ie/wp/wp-content/uploads/2015/06/Keating-PP-PWS-.pdf>

FIGURE B.3

Example of a Google Scholar advanced search entry

Advanced search

Find articles

with all of the words

with the exact phrase

with at least one of the words

without the words

where my words occur

Return articles authored by

Return articles published in

Return articles dated between

anywhere in the article

in the title of the article

e.g., "PJ Hayes" or McCarthy

e.g., J Biol Chem or Nature

e.g., 1996

Step 4: Searching for impact evaluations listed by other clearinghouses

As a supplementary step, you might want to see if other clearinghouses provide details of other impact evaluations of your programme of interest. For example, you might want to visit the following clearinghouse websites:

- Blueprints for Healthy Youth Development⁹
- California Evidence-Based Clearinghouse for Child Welfare.¹⁰

Step 5: Extracting information from the impact evaluations identified

Once you have found and accessed relevant impact evaluations, you can read the 'Methodology' and 'Measures' sections of each paper to identify the measures used by the authors to assess their outcomes of interest. For each study, try to identify the outcomes assessed and the measures used. Several measures have more than one version, so make sure you have the full name of the measures used and the reference to the original paper.

To organise your findings, you may consider creating a table like the one below. Once complete, you can then consider which of the measures extracted might be most relevant for your own evaluation, taking into consideration the recommendations made in the guidance section of this report (see chapter 5).

⁹ See: <http://www.blueprintsprograms.com/programs>

¹⁰ See: <https://www.cebc4cw.org/>

TABLE B.1

Intended outcomes and measures used in the impact evaluation of Parents Plus Parenting when Separated, according to study 1, Keating et al. (2015)

| | |
|--------------------------|--|
| Intended outcomes | Child outcomes Improved child behaviour Interparental (or couple) outcomes Reduced interparental conflict and improved co-parenting and conflict management skills Parent outcomes Improved perception of parent–child relationship quality Improved psychological adjustment |
| Measures used | Child outcome measures Strengths and Difficulties Questionnaire (SDQ parent-report; Goodman, 2001), specifically, the Total Difficulties Scale Couple outcome measures Quality of Co-parental Communications Scale (QCCS; Ahrons, 1981), specifically, the Interparental Conflict subscale Parent outcome measures Kansas Parental Satisfaction Scale (KPS self-report; James et al., 1985) Mental Health Inventory (MHI-5; Berwick et al., 1991) |

Appendix C: Methodology

This report is based on a rapid evidence assessment of a selected number of measures relevant to the evaluation of programmes targeting interparental relationships. Although a systematic review would have been the most rigorous approach, we did not have sufficient time or resources for that, and so we conducted a rapid evidence assessment which was comprised of the following steps:

1. Selecting the relevant measures for inclusion in the review
2. Developing our measurement assessment criteria
3. Conducting a rapid review of the literature published on our selected measures and applying the assessment criteria to each measure

Each of the three components is discussed in further detail below.

Advisory group involvement

As part of this project, a group of advisory members was set up consisting of subject-matter experts including academics, practitioners and providers, who contributed with valuable input and quality assurance throughout the study design and write-up. We also assembled a small group of psychometric experts, who helped us develop our measurement assessment criteria, troubleshoot through issues regarding the application of the criteria, and ensure that we were applying our criteria consistently. Finally, we reached out to a group of local authority representatives in order to ensure that the messages of our report were well suited to our target audience. For more details on the advisory group, please see appendix D.

Step 1: Selecting the relevant measures

Selecting outcomes of interest

To identify the most suitable measures for local areas to use when assessing progress for individual families taking part in programmes targeting the interparental relationship, we first had to define our main outcomes of interest. Although parental conflict is known to impact on a wide range of outcomes for both parents and children, given time and resource constraints, we had to restrict the scope of this review to only a core set of outcomes.

Child outcomes:

- **internalising behaviours**, including symptoms of anxiety or depression
- **externalising behaviours**, including antisocial or aggressive behaviours.

Interparental (or couple) relationship outcomes:

- **relationship quality**, including factors like interparental or couple satisfaction, commitment, communication, respect, shared interests and dyadic consensus
- **parental conflict**, specifically the frequency and intensity of conflict, conflict resolution skills, and children's perception of and adjustment to the conflict between parents
- **co-parenting practices**, including parents' ability to cooperate and communicate when performing parenting duties.

Outcomes not within scope for this review included broader child outcomes (such as social and interpersonal relationship skills, academic performance and physical health) and parent outcomes (such as mental health, parenting practices and parental self-efficacy), as well as family outcomes such as the quality of parent–child and sibling relationships. Moreover, because the focus of this report was on non-abusive conflict between parents, we did not include measures assessing domestic abuse.

For more details on the rationale behind our choice of outcomes, please see section 2.3.

Identifying a longlist of measures

Once our outcomes of interest had been selected, we sought to identify a longlist of suitable measures by collating measures that were:

- used in impact evaluations of programmes targeting interparental relationships, which were identified in our 2017 review on parental conflict in the context of poverty,¹¹ but mainly drawn from our comprehensive review published in 2016 in collaboration with Professor Gordon Harold¹²
- used in evaluations of the eight face-to-face interventions currently being trialled in the UK as part of DWP’s Reducing Parental Conflict Programme¹³
- used in DWP’s own evaluation of the national programme.

We then also contacted our groups of expert advisors and local authority leaders, and asked them to identify relevant measures which we may have missed and would benefit from inclusion in this review.

Altogether we identified 233 measures; 147 assessing interparental (or couple) relationship outcomes and 86 assessing child outcomes.

Filtering and selecting the relevant measures

In order to determine which measures should be excluded from the study and which should be considered for an in-depth review, the following eligibility criteria were applied to our longlist of 233 measures.

Inclusion criteria – a measure was included if it:

- could be used to assess progress for individuals taking part in programmes targeting the couple/interparental relationship
- was designed to assess at least one of our predefined outcomes of interest
- was quantitative (for example not observational or interview-based)
- was widely used (for example used in at least five evaluations)
- was used outside of the specific programme for which it had been designed
- was developed, updated or used in an economically developed country in the last 10 years.

¹¹ Acquah, D., Sellers, R., Stock, L., & Harold, G. (2017). *Inter-parental conflict and outcomes for children in the contexts of poverty and economic pressure*. London: Early Intervention Foundation. <https://www.eif.org.uk/report/interparental-conflict-and-outcomes-for-children-in-the-contexts-of-poverty-and-economic-pressure>

¹² Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance interparental relationships and improve outcomes for children*. Early Intervention Foundation: London. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>

¹³ See: <https://www.eif.org.uk/files/pdf/cg-rpc-4-3-face-to-face-support-interventions.pdf>

Exclusion criteria – a measure was excluded if it:

- was a duplicate
- was primarily used for assessing outcomes other than those defined as our outcomes of interest (for example exclusively focused on assessing violent and/or abusive relationships)
- was not available in English
- was solely used as a screening or diagnostic tool
- was unpublished, had been published outside of a peer-reviewed journal, or had no full-text article available that either described or psychometrically evaluated the measure.

Despite our reliance on the eligibility criteria listed above, we were required to conduct a second filtering process because we still had a total of 56 measures, which would have been too many to include in this review. We therefore then adopted a flexible approach to the selection of measures, by prioritising those which we knew were commonly used by local areas. We also tried to select measures which assess an adequate range of outcomes and target the broadest age groups. For instance, if we had three measures assessing symptoms of anxiety, we endeavoured to select the measure which was most commonly used by local areas and which assessed the broadest age group.

Search results

As illustrated in figure C.1, we included a total of 18 measures from an initial longlist of 233. Table C.1 provides a breakdown of the 18 measures selected for inclusion in this review, with 11 assessing interparental (or couple) relationship outcomes and 7 assessing child outcomes.

FIGURE C.1

Flow diagram of measurement selection

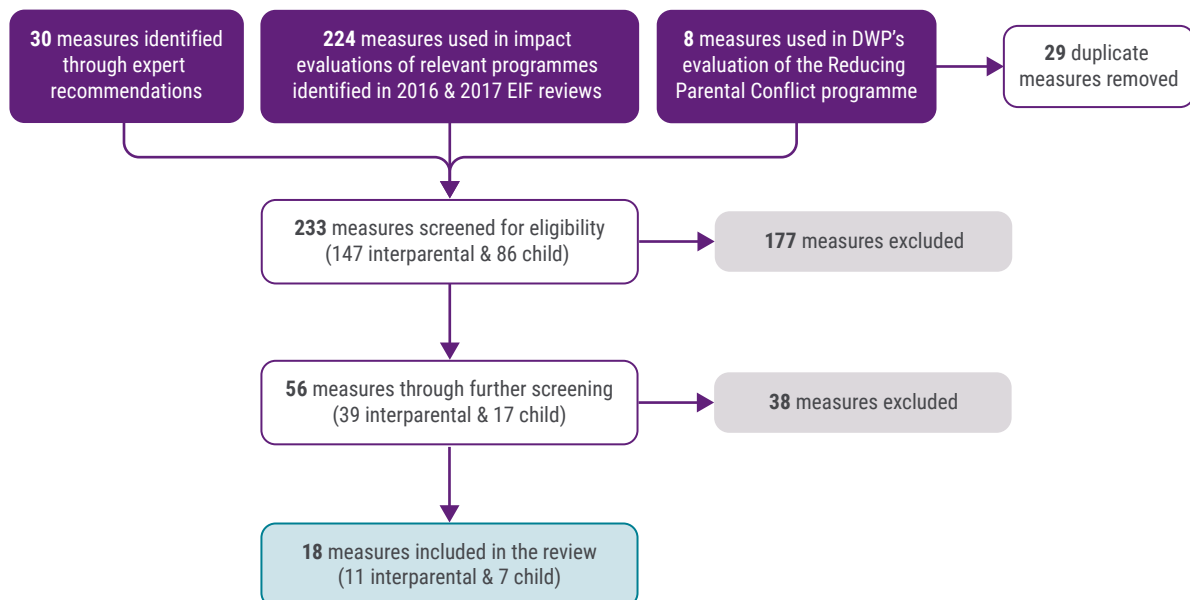


TABLE C.1

Measures selected for inclusion in this review

| Measures | Total |
|--|-----------|
| Child outcome measures | 7 |
| Internalising and externalising behaviours | 2 |
| Child Behaviour Checklist (CBCL/6–18) | |
| Strengths and Difficulties Questionnaire (SDQ) | |
| Internalising behaviours | 3 |
| Mood and Feelings Questionnaire (MFQ) | |
| Revised Child Anxiety and Depression Scale (RCADS) | |
| Screen for Child Anxiety Related Emotional Disorders (SCARED) | |
| Externalising behaviours | 2 |
| Eyberg Child Behavior Inventory (ECBI) | |
| Parent Daily Report (PDR) | |
| Interparental (or couple) relationship outcome measures | 11 |
| Relationship quality | 6 |
| Couples Satisfaction Index (CSI-16) | |
| Dyadic Adjustment Scale (DAS-32) | |
| Dyadic Adjustment Scale (DAS-7) | |
| Golombok Rust Inventory of Marital State (GRIMS) | |
| Relationship Quality Index (RQI) | |
| Marital Adjustment Test (MAT) | |
| Parental conflict | 2 |
| Children's Perception of Interparental Conflict Scale (CPIC) | |
| O'Leary Porter Scale (OPS) | |
| Co-parenting | 3 |
| Parenting Alliance Measure (PAM) | |
| Parent Problem Checklist (PPC) | |
| Quality of Co-parental Communication Scale (QCCS) | |

Step 2: Developing our measurement assessment criteria

Once the relevant measures had been selected for inclusion in this review, the research team focused on the development of a measurement assessment criteria, which would be used to assess the validity, reliability and usability of each measure. Based on a review of the measurement guidelines used by eight different clearinghouses, we identified the following selection of important psychometric and implementation criteria.

Psychometric criteria:

- internal consistency
- test-retest reliability
- validity
- sensitivity to change.

Implementation criteria:

- brevity
- availability
- ease of scoring.

Since the aim of this project was to identify the most suitable measures for UK local areas to use when assessing the impact of programmes targeting the interparental (or couple) relationship, it was important for us to let our readers know whether the selected measures had been previously used with a UK population. Consequently, we decided to include one further implementation criterion:

- Used in the UK

Once we had identified the relevant criteria to include as part of our measurement assessment, we then had to decide on the specifics of each criteria. This was a complex exercise to complete, especially when considering that each clearinghouse used a slightly different psychometric cut-off score. To address this, we first created a table listing all the proposed cut-offs from each clearinghouse. We then compared the details of each cut-off score and created our own criteria by striking a balance between these different ones – aiming for a reasonable and robust cut-off value.

As soon as our provisional criteria had been finalised, we asked for feedback from our group of psychometric experts. The comments we received from our panel of psychometric experts were then discussed at an internal meeting, after which some of the criteria were reviewed in accordance with the feedback. The final criteria listed in table C.2 was used in the final step of our methodology, to conduct an in-depth assessment of the measures selected for inclusion in this review.

TABLE C.2

Measurement assessment criteria

| Features | Description | Criteria |
|---|--|---|
| Psychometric features | | |
| Internal consistency | The degree to which items designed to measure the same outcome relate to one another. | Cronbach's alpha coefficients ≥ 0.60 . |
| Test-retest reliability | The extent to which the outcomes of an assessment are stable over time. Assuming no important intervening events, a person's score on a measure taken multiple times should be correlated. | Test-retest scores (for example ICC) ≥ 0.70 for short periods of time (1–4 weeks). |
| Validity Criterion validity AND/OR Construct validity AND/OR Concurrent validity | Criterion validity: The extent to which scores on a particular measure are related to a 'gold standard'. Construct validity: The extent to which scores on a particular measure relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the measured concepts. Concurrent validity: The extent to which scores on a new measure are related to scores from a criterion measure administered at the same time. | Criterion validity is reported, with a convincing argument that the gold standard is 'gold' and the correlation with that gold standard (typically AUC) is ≥ 0.70 AND/OR Pearson correlation tested against similar measures $\geq 0.70^*$ AND/OR Confirmatory Factor Analysis shows strong results (CFI ≥ 0.95 and RMSEA < 0.05). |
| Sensitivity to change | The ability of a measure to detect important changes over time (post-intervention), even if these changes are small. | The measure has been used and found to identify change over time in two impact evaluations (for example pre-posts/QEDs/RCTs) of relevant interventions conducted over a short period of time. |

| Implementation features | | |
|-------------------------|---|--|
| Brevity | The time taken to complete a measure and/or the number of items in a measure. | The measure takes ≤ 15 minutes to complete AND/OR It comprises of ≤ 30 items. |
| Availability | The extent to which a measure is freely available. | The measure is free to use. |
| Ease of scoring | The extent to which a measure is easy to score and interpret. | The measure has simple scoring instructions AND It does not need to be scored/interpreted by someone with specific training or qualifications. |
| Used in the UK | The extent to which the measure has been used in the UK setting. | The measure has been used in at least two UK studies published in peer-reviewed journals by different research teams. |

Note: *If < 0.70, case-by-case decisions are made.

Step 3: Applying our measurement assessment criteria

Developing a search strategy

Once we had selected our measures of interest and developed a measurement assessment criteria, we then had to identify relevant papers describing the design, development and psychometric testing of these measures. Given time and resource constraints, we were unable to conduct a systematic review of all the literature published on our selected measures. Considering the vast literature available on each of these measures, we conducted a rapid review of the evidence and developed the following search strategy to ensure that we identified the most relevant studies. This strategy also helped to certify that each member of the research team was using the same approach.

Search strategy

- Our starting point was always the original paper describing the development and psychometric testing of the version of the measure we were assessing.
- When the measure's manual was freely available, we used it to extract relevant psychometric and implementation information.
- Then, where available, we used the reference list provided by the official website. When possible, we filtered the references to find information only on the version, in English, that we were assessing. We always excluded studies based on the translated versions of the measure. When it was not possible to filter the list, we read each abstract to decide whether to include a study or not. We only accepted studies published in peer-reviewed journals and conducted with a sample size of more than 100 people. An exception was made for studies reporting evidence of sensitivity to change and the criterion 'used in the UK'. In line with EIF's Programme Assessment procedure, we accepted studies conducted with a sample of at least 20 people.
- Next, we searched for systematic reviews and meta-analyses assessing the measure of interest. Here too we only included reviews assessing the relevant version of the measure. When the authors did not report the version of the measure used in their review, and it was not possible for us to infer this information from the reference list provided, we excluded the systematic review/meta-analysis from our study.

- We supplemented the above steps with targeted searches of Google Scholar and PubMed using predefined search terms to fill identified gaps related to specific criteria. In the case of PubMed, we used the search filters proposed by Terwee and colleagues.¹⁴
- Finally, when we found consistent references to the manual on specific psychometric information, and the manual was not freely available, we contacted the developers and asked if they could share with us that specific piece of evidence. While in most cases the developers agreed to share the information reported in the manual, in some cases we were not able to obtain the evidence we were looking for.

Applying the criteria

To apply our assessment criteria, we extracted relevant information from each of the studies identified via our aforementioned search strategy, and reviewed the gathered evidence to establish whether the measure under review had met each of our criterion. To score each criterion, we developed the following rating system:

| | |
|---|--|
| ✓ | We found clear evidence that the selected version of the measure had met the criterion |
| × | We were unable to find evidence showing that the selected version of the measure had met the criterion |
| ? | We found insufficient evidence to determine if the criterion had been met |

To ensure inter-rater reliability, 20% of the measures were assessed independently by the two members of the research team, and their assessments compared. The percentage agreement was 81%, suggesting that the search strategy, study selection process and application of the criteria were sufficiently standardised to yield similar results.

Applying the psychometric criteria

In several cases, the evidence related to a specific psychometric criterion reported values that were very close to our specified cut-offs but did not meet our criteria. In other cases, the selected studies reported relevant information that had not been included in our criteria. Finally, in some cases the selected studies reported conflicting evidence, with some values meeting our criteria and others not. To quality assure our application of the criteria, each unclear case was discussed at internal meetings, during which a consensus decision was reached. Particularly difficult cases were also discussed with our panel of psychometric experts, either via email correspondence or during our designated psychometric meeting.

An illustrating example is the process through which we developed and applied our validity criteria. Initially, our validity criteria referred only to criterion validity (the extent to which scores on a particular measure are related to a 'gold standard') and construct validity (the extent to which scores on a particular measure relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the measured concepts). When we started applying such criteria, however, we observed that numerous studies reported Pearson Correlation coefficients as evidence of concurrent validity (the extent to which scores on a new measure are related to scores from a criterion measure administered at the same time).

¹⁴ Terwee, C.B., Jansma, E.P., Riphagen, I.I., & de Vet, H.C. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*, 18(8), 1115–1123.

To decide whether to change our validity criteria to include also concurrent validity, we discussed this possibility both internally and with our panel of psychometric experts. We finally opted for including concurrent validity within our broader validity criteria. Since Pearson correlation coefficients depend on the measure chosen to test concurrent validity, we decided to examine case by case every Pearson correlation coefficient below 0.70, which was considered to be an acceptable value. For each case, we scored the measure by examining the similarity (in terms of outcomes assessed, length of the measure, and who completed it) between the measure we were assessing, and the measure used to test its validity. The most complex cases were discussed with our panel of psychometric experts.

Applying the implementation criteria

In some cases, the gathered evidence was not sufficient to establish whether the measure under review was sufficiently brief, freely available and easy to score. Where possible, we contacted the developer(s) of the measure in question, to ask for further information in order to score our criteria.

For the brevity criterion, we decided to prioritise completion time over the number of items. Consequently, if a measure had more than 30 items but there was sufficient evidence to suggest that the completion time was less than 15 minutes, the criterion was considered met.

Quality assurance

Once the criteria had been applied to the majority of measures, our three psychometric experts were invited to a one-day quality assurance meeting during which we discussed specific problems with applying our criteria, and reached agreement on how to score particular measures. The assessments of the measures we discussed were changed according to the decisions made at this meeting, and new measures were assessed accordingly. All other issues encountered were discussed with the panel via email correspondence.

Creating measurement reports

For each of the measures selected, we created a report with a summary of how each measure met our assessment criteria. We also described the main features of the measure, including the following details:

- name of measure and acronym
- developer(s) and publication year
- versions available (if applicable)
- outcome(s) assessed
- target population
- scales and subscales (if applicable)
- mode of administration
- example item
- response format
- strengths and limitations
- link and contact details
- copyright information
- key reference(s).

Methodological strengths and limitations

Strengths

- Although this is not a full systematic review, the methodology used to identify the relevant measures and apply the measurement assessment criteria is clear. We are confident that, through the search strategy used, we have identified most of the relevant papers describing the design, development and psychometric testing of each measure.
- The methodology we used to apply the assessment criteria demonstrated good inter-rater reliability, suggesting that the search strategy, study selection process and application of the criteria were sufficiently standardised to yield similar results.
- Our process of quality assurance was strengthened by input from psychometric experts.

Limitations

- This guide does not include an exhaustive list of measures.
- Given our methods for identifying a longlist of measures and filtering this longlist to a select number of measures, it is possible that we have missed out important measures assessing our predefined outcomes of interest.
- Due to time and resource constraints, we were not able to assess all the different versions of the selected measures. This may, for example, mean that certain age groups (for instance 0–4 years) are underrepresented in this review, as we opted to select the versions of measures assessing the outcomes of children with the broadest age range, which was typically over 4–6 years of age.
- The methodological approach used in this review did not involve an exhaustive search of the literature. There is therefore a risk that we have missed relevant evidence when applying our measurement assessment criteria, meaning that our assessment of each measure may be incomplete. This is likely to be a particular issue for criteria items marked as 'X' (indicating that the values did not meet the criterion) and/or '?' (indicating that there was insufficient evidence to determine whether the criterion had been met).

Appendix D: Advisory members

As part of this project, a group of advisory members was set-up consisting of subject-matter experts including academics, practitioners and providers, who contributed with valuable input and quality assurance throughout the design and write-up of this guide. We also assembled a small group of psychometric experts, who helped us develop our measurement assessment criteria, troubleshoot through issues regarding the application of the criteria, and ensure that we were applying our criteria consistently. Finally, we reached out to a group of local authority representatives, in order to ensure that the messages of our report were well suited to our target audience.

We are grateful for the contribution of all our advisory members (listed here alphabetically, by surname).

Subject-matter experts

- Professor Jane Barlow, University of Oxford
- Professor Jacqueline Barnes, Birkbeck University
- Dr Shannon Hirst, OnePlusOne
- Dr Emma Howarth, University of East London
- Jenny Porter, Marriage Care
- Ailsa Redhouse, Department for Work and Pensions
- Honor Rhodes, Tavistock Relationships
- Dr Camilla Rosan, Anna Freud National Centre for Children and Families
- Luke Staniland, Department for Work and Pensions

Psychometric experts

- Professor Jacqueline Barnes, Birkbeck University
- Dr Ela Polek, Tavistock Relationships
- Dr Silia Vitoratou, King's College London

Local authority representatives

- Helen Armstrong, Regional Integration Lead for the North West, Reducing Parental Conflict Programme
- Bridie Collins, Westminster
- Tania Davies, Plymouth and Torbay
- Ruth Fennemore, Hertfordshire
- Tracy Lysons, Blackburn with Darwen Borough Council
- Emily Nickson-Williams, Rochdale

The following experts, while not formally part of our advisory group, were also asked to contribute with their suggestions of measures: Professor Susan Ayers (City University London), Dr Crispin Day (King's College London/Maudsley NHS Foundation Trust), Dr Erica Hepper (University of Surrey) and Professor Paul Ramchandani (University of Cambridge).

References

1. Introduction

- Acquah, D., Sellers, R., Stock, L., & Harold, G. (2017). *Inter-parental conflict and outcomes for children in the contexts of poverty and economic pressure*. London: Early Intervention Foundation. <https://www.eif.org.uk/report/interparental-conflict-and-outcomes-for-children-in-the-contexts-of-poverty-and-economic-pressure>
- Department for Children, Schools and Families. (2010). *Support for all: The Families and Relationships Green Paper*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/329448/Support_for_All.pdf
- Department for Work and Pensions. (2017a). *Improving Lives: Helping Workless Families*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/621364/improving-lives-helping-workless-families-web-version.pdf
- Department for Work and Pensions. (2017b). *Improving Lives: Helping Workless Families Analysis and Research Pack*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696368/improving-lives-helping-workless-families-web-version.pdf
- Department of Health and Social Care and Department for Education. (2017). *Transforming Children and Young People's Mental Health Provision: A Green Paper*. London: HM Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/664855/Transforming_children_and_young_people_s_mental_health_provision.pdf
- Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance interparental relationships and improve outcomes for children*. Early Intervention Foundation: London. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>
- Harold, G., & Sellers, R. (2018). Annual Research Review: Interparental conflict and youth psychopathology – an evidence review and practice focused update. *Journal of Child Psychology and Psychiatry*, 59(4), 374-402. <http://sro.sussex.ac.uk/id/eprint/73500/>
- Martin, J., McBride, T., Brims, L., Doubell, L. Pote, I., & Clarke, A. (2018). *Evaluating early intervention programmes: Six common pitfalls, and how to avoid them*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/evaluating-early-intervention-programmes-six-common-pitfalls-and-how-to-avoid-them>
- Taylor, S., Drayton, E., & McBride, T. (2019). *Evaluating early help: A guide to evaluation of complex local early help systems*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/evaluating-early-help-a-guide-to-evaluation-of-complex-local-early-help-systems>

2. A brief summary of the evidence on parental conflict

- Conger, R.D., Ge, X., Elder, G.H., Lorenz, F.O., & Simons, R.L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, 65(2), 541–561.
- Cummings, E., & Davies, P. (2002). Effects of marital conflict on children: Recent advances and emerging themes in process-oriented research. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(1), 31–63.
- Davies, P.T., & Cummings, E.M. (1994). Marital conflict and child adjustment: An emotional security hypothesis. *Psychological Bulletin*, 116(3), 387–411.
- Erel, O., & Burman B. (1995). Interrelatedness of marital relations and parent-child relations: A meta-analytic review. *Psychological Bulletin*, 118(1), 108–132.
- Grych, J.H., & Fincham, F.D. (1990). Marital conflict and children's adjustment: A cognitive-contextual framework. *Psychological Bulletin*, 108(2), 267–290.
- Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance interparental relationships and improve outcomes for children*. Early Intervention Foundation: London. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>
- Harold, G.T., & Conger, R.D. (1997). Marital conflict and adolescent distress: The role of adolescent awareness. *Child Development*, 68(2), 333–350.

3. Why measure?

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., & de Vet, H.C.W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63, 737–745.

Souza, A.C.D., Alexandre, N.M.C, and Guirardello, E.D.B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, 26(3), 649–659.

4. Measures

Acquah, D., Sellers, R., Stock, L., & Harold, G. (2017). *Inter-parental conflict and outcomes for children in the contexts of poverty and economic pressure*. London: Early Intervention Foundation. <https://www.eif.org.uk/report/interparental-conflict-and-outcomes-for-children-in-the-contexts-of-poverty-and-economic-pressure>

Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance interparental relationships and improve outcomes for children*. Early Intervention Foundation: London. <https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children>

5. Guidance on selecting and using outcome measures

Asmussen, K., Brims, L., & McBride, T. (2019). *10 steps for evaluation success*. London: Early Intervention Foundation. <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>

Institute for Employment Studies. (2001). *Guide to Measuring Soft Outcomes and Distance Travelled*. <http://webarchive.nationalarchives.gov.uk/20090902143103/http://readingroom.lsc.gov.uk/lsc/SouthEast/distance1.pdf>

Karekla, M., Kasinopoulos, O., Neto, D.D., Ebert, D.D., Van Daele, T., Nordgreen, T., Höfer, S., Oeverland, S., & Jensen, K.L. (2019). Best practices and recommendations for digital interventions to improve engagement and adherence in chronic illness sufferers. *European Psychologist*, 24(1), 49–67.

Medical Research Council [MRC] Regulatory Support Centre and NHS Health Research Authority. (2018, September). Do I need NHS REC approval?. <http://www.hra-decisiontools.org.uk/ethics/>

Mushquash, C.J., & Bova, D.L. (2007). Cross-cultural assessment and measurement issues. *Journal on Developmental Disabilities*, 13(1), 53–65.

NHS Health Research Authority. (2018, June 13). HRA guidance on the General Data Protection Regulation. <https://www.hra.nhs.uk/hra-guidance-general-data-protection-regulation/>

NHS Health Research Authority. (n.d.). Approvals and amendments. <https://www.hra.nhs.uk/approvals-amendments/>

Yardley, L., Spring, B.J., Riper, H., Morrison, L.G., Crane, D.H., Curtis, K., Merchant, G.C., Naughton, F., & Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833–842.